

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds

(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083a-31

Perfect score: 49

Sequence: 1 KWEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 366760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Genesep16Dec04:\*  
2: genesep1980s:\*  
3: genesep1990s:\*  
4: genesep2000s:\*  
5: genesep2001s:\*  
6: genesep2002s:\*  
7: genesep2003as:\*  
8: genesep2003bs:\*  
9: genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	9	4 AAE08294	Aae08294 Human str
2	49	100.0	9	4 AAE08236	Aae08236 Human str
3	49	100.0	9	8 ADR68850	Adr68850 Human str
4	49	100.0	9	8 ADR68792	Adr68792 Human str
5	49	100.0	9	8 ADR68870	Adr68870 Human str
6	49	100.0	97	6 ADA05740	Ada05740 Human NOV
7	49	100.0	97	8 ADN62904	Adn62904 Human NOV
8	49	100.0	136	4 ABG23378	Abg23378 Novel hum
9	49	100.0	144	8 ADI39727	Adi39727 Stratum c
10	49	100.0	144	8 ADI37151	Adi37151 Stratum c
11	49	100.0	181	6 ADA05738	Ada05738 Human NOV
12	49	100.0	181	8 ADN62902	Adn62902 Human NOV
13	49	100.0	198	6 ADA05736	Ada05736 Human NOV
14	49	100.0	198	8 ADN62900	Adn62900 Human NOV
15	49	100.0	224	6 ADA05744	Ada05744 Human NOV
16	49	100.0	224	8 ADN62908	Adn62908 Human NOV
17	49	100.0	225	4 AAB98502	Aab98502 Human NOV
18	49	100.0	247	6 ADA05742	Ada05742 Human NOV
19	49	100.0	247	8 ADN62906	Adn62906 Human NOV
20	49	100.0	250	6 ADA05732	Ada05732 Human NOV
21	49	100.0	250	8 ADN62896	Adn62896 Human NOV
22	49	100.0	252	6 ADA05734	Ada05734 Human NOV
23	49	100.0	252	8 ADN62898	Adn62898 Human NOV
24	49	100.0	253	2 AAR67888	Aar67888 Human str
25	49	100.0	253	2 AAW05383	Aaw05383 Human amy

26	49	100.0	253	5 ABB84421	Abb84421 Human SCC
27	49	100.0	253	5 ABB84406	Abb84406 Human SCC
28	49	100.0	253	5 AAU82740	Aau82740 Amino aci
29	49	100.0	253	6 ABU07440	Abu07440 Protein d
30	49	100.0	253	6 ABU07471	Abu07471 Protein d
31	49	100.0	253	6 ABR58471	AbR58471 Human str
32	49	100.0	253	7 ADB80484	AdB80484 Ovarian c
33	49	100.0	253	7 ADJ68833	AdJ68833 Human hea
34	49	100.0	253	7 ADN39180	Adn39180 Cancer/an
35	49	100.0	253	8 ADL06515	AdL06515 Human tum
36	49	100.0	253	8 ADN04182	AdN04182 Antipsori
37	49	100.0	253	8 ADR72880	AdR72880 Human ova
38	49	100.0	257	3 AAB21326	Aab21326 Human HSC
39	40	81.6	226	5 ABB84422	Abb84422 Rat SCCE
40	39	79.6	9	4 AAE08326	Aae08326 Human str
41	39	79.6	9	8 ADR68883	Adr68883 Human str
42	37	75.5	9	4 AAE08298	Aae08298 Human str
43	37	75.5	9	8 ADR68854	Adr68854 Human str
44	37	75.5	243	5 ABB84419	Abb84419 Bovine SC
45	37	75.5	334	4 AAU37656	Aau37656 Streptoco

## ALIGNMENTS

RESULT 1  
AAE08294  
ID AAE08294 standard; peptide; 9 AA.  
XX AAE08294;  
XX  
DT 01-NOV-2001 (first entry)  
XX  
DE Human stratum corneum chymotrypsin enzyme peptide #59 (residues 72-80).  
XX  
KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;  
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;  
KW antisense therapy; malignant hyperplasia.  
XX  
OS Homo sapiens.  
XX  
PN WO200159158-A1.  
XX  
PD 16-AUG-2001.  
XX  
PF 07-FEB-2001; 2001WO-US003977.  
XX  
PR 11-FEB-2000; 2000US-00502600.  
XX  
PA (UTAR-) UNIV ARKANSAS.  
XX  
PI O'brien TJ;  
XX  
DR WPI; 2001-514676/56.  
XX  
PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.  
XX  
PS Disclosure; Page 115; 127pp; English.  
XX  
CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

SQ Sequence 9 AA;

Query Match 100.0%; Score 49; DB 4; Length 9;

```
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 1 KMNEYTVHL 9

RESULT 2
AAE08236
ID AAE08236 standard; peptide; 9 AA.
XX AC AAE08236;
XX DT 01-NOV-2001 (first entry)
XX DE Human stratum corneum chymotrypsin enzyme peptide #1 (residues 72-80).
XX KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
XX KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
XX KW antisense therapy; malignant hyperplasia.
XX OS Homo sapiens.
XX PN WO200159158-A1.
XX PD 16-AUG-2001.
XX PF 07-FEB-2001; 2001WO-US003977.
XX PR 11-FEB-2000; 2000US-00502600.
XX PA (UYAR-) UNIV ARKANSAS.
XX PI O'brien TJ;
XX DR WPI; 2001-514676/56.
XX PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin
XX PT enzyme.
XX PS Claim 25; Page 102; 127pp; English.
XX CC The invention relates to diagnosing cancer especially ovarian cancer, by
XX CC screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are
XX CC considered to be an integral part of tumour growth and metastasis, and
XX CC therefore, markers indicative of their presence or absence are useful for
XX CC the diagnosis of cancer. The method is useful for diagnosing cancer,
XX CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used
XX CC to treat a cancer selected from ovarian, breast, lung, colon, prostate
XX CC and other cancers in which SCCE is overexpressed. The present sequence is
XX CC human SCCE peptide
XX SQ Sequence 9 AA;

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 1 KMNEYTVHL 9

RESULT 3
ADR68850
ID ADR68850 standard; peptide; 9 AA.
XX AC ADR68850;
XX DT 02-DEC-2004 (first entry)
XX DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:89.
XX
```

```
XX serine protease; stratum corneum chymotrytic enzyme; SCCE;
KW immune response; ovarian cancer; lung cancer; prostate cancer;
KW pancreatic cancer; colon cancer.
XX OS Homo sapiens.
XX PN WO2004075723-A2.
XX PD 10-SEP-2004.
XX PF 20-FEB-2004; 2004WO-US005134.
XX PR 21-FEB-2003; 2003US-00372521.
XX PA (UYAR-) UNIV ARKANSAS.
XX PI O'brien TJ, Cannon MJ, Santin A;
XX DR WPI; 2004-653294/63.
XX PT Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for
XX PT vaccinating an individual against SCCE, and in monitoring and developing
XX PT immunotherapies for ovarian and other malignancies.
XX PS Disclosure; SEQ ID NO 89; 117pp; English.
XX CC The invention relates to a novel method for vaccinating an individual
XX CC against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating
XX CC an individual with a SCCE peptide, which elicits an immune response in
XX CC the individual. A peptide of the invention acts as a stratum corneum
XX CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating
XX CC an individual against SCCE, particularly an individual having, suspected
XX CC or at risk of getting ovarian, lung, prostate, pancreatic or colon
XX CC cancer. The oligonucleotide is useful for treating a neoplastic state in
XX CC an individual, such as ovarian, breast, lung, colon, prostate, or
XX CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The
XX CC peptides are also useful in the monitoring and development of
XX CC immunotherapies for ovarian and other malignancies. The present sequence
XX CC represents a peptide fragment of serine protease SCCE (stratum corneum
XX CC chymotrytic enzyme).
XX SQ Sequence 9 AA;

Query Match 100.0%; Score 49; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 1 KMNEYTVHL 9

RESULT 4
ADR68792
ID ADR68792 standard; peptide; 9 AA.
XX AC ADR68792;
XX DT 02-DEC-2004 (first entry)
XX DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:31.
XX KW serine protease; stratum corneum chymotrytic enzyme; SCCE;
KW immune response; ovarian cancer; lung cancer; prostate cancer;
KW pancreatic cancer; colon cancer.
XX OS Homo sapiens.
XX PN WO2004075723-A2.
XX PD 10-SEP-2004.
XX
```



PF 20-FEB-2004; 2004WO-US005134.  
 XX  
 PR 21-FEB-2003; 2003US-00372521.  
 XX  
 PA (UYAR-) UNIV ARKANSAS.  
 XX  
 PI O'Brien TJ, Cannon MJ, Santin A;  
 XX WPI; 2004-653294/63.  
 DR  
 XX Use of stratum corneum chymotryptic enzyme (SCCE) peptides, for  
 PT vaccinating an individual against SCCE, and in monitoring and developing  
 PT immunotherapies for ovarian and other malignancies.  
 XX  
 PS Claim 5; SEQ ID NO 31; 117pp; English.  
 XX  
 CC The invention relates to a novel method for vaccinating an individual  
 CC against stratum corneum chymotryptic enzyme (SCCE), comprising inoculating  
 CC an individual with a SCCE peptide, which elicits an immune response in  
 CC the individual. A peptide of the invention acts as a stratum corneum  
 CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
 CC an individual against SCCE, particularly an individual having, suspected  
 CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
 CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
 CC an individual, such as ovarian, breast, lung, colon, prostate, or  
 CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
 CC peptides are also useful in the monitoring and development of  
 CC immunotherapies for ovarian and other malignancies. The present sequence  
 CC represents a peptide fragment of serine protease SCCE (stratum corneum  
 CC chymotrytic enzyme).  
 XX  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 49; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KNNEYTVHL 9  
 Db |||||  
 1 KNNEYTVHL 9  
 RESULT 6  
 ADA05740  
 ID ADA05740 standard; protein; 97 AA.  
 XX  
 AC ADA05740;  
 XX  
 DT 06-NOV-2003 (first entry)  
 DE Human NOVI8e protein SEQ ID NO:100.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003029424-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 02-OCT-2002; 2002WO-US031373.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.

PF 20-FEB-2004; 2004WO-US005134.  
 XX  
 PR 21-FEB-2003; 2003US-00372521.  
 XX  
 PA (UYAR-) UNIV ARKANSAS.  
 XX  
 PI O'Brien TJ, Cannon MJ, Santin A;  
 XX WPI; 2004-653294/63.  
 DR  
 XX Use of stratum corneum chymotryptic enzyme (SCCE) peptides, for  
 PT vaccinating an individual against SCCE, and in monitoring and developing  
 PT immunotherapies for ovarian and other malignancies.  
 XX  
 PS Claim 5; SEQ ID NO 31; 117pp; English.  
 XX  
 CC The invention relates to a novel method for vaccinating an individual  
 CC against stratum corneum chymotryptic enzyme (SCCE), comprising inoculating  
 CC an individual with a SCCE peptide, which elicits an immune response in  
 CC the individual. A peptide of the invention acts as a stratum corneum  
 CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
 CC an individual against SCCE, particularly an individual having, suspected  
 CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
 CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
 CC an individual, such as ovarian, breast, lung, colon, prostate, or  
 CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
 CC peptides are also useful in the monitoring and development of  
 CC immunotherapies for ovarian and other malignancies. The present sequence  
 CC represents a peptide fragment of serine protease SCCE (stratum corneum  
 CC chymotrytic enzyme).  
 XX  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 49; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KNNEYTVHL 9  
 Db |||||  
 1 KNNEYTVHL 9  
 RESULT 5  
 ADR68870  
 ID ADR68870 standard; peptide; 9 AA.  
 XX  
 AC ADR68870;  
 XX  
 DT 02-DEC-2004 (first entry)  
 DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:109.  
 XX  
 KW serine protease; stratum corneum chymotrytic enzyme; SCCE;  
 KW immune response; ovarian cancer; lung cancer; prostate cancer;  
 KW pancreatic cancer; colon cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004075723-A2.  
 XX  
 PD 10-SEP-2004.  
 XX  
 PF 20-FEB-2004; 2004WO-US005134.  
 XX  
 PR 21-FEB-2003; 2003US-00372521.  
 XX  
 PA (UYAR-) UNIV ARKANSAS.  
 XX  
 PI O'Brien TJ, Cannon MJ, Santin A;  
 XX WPI; 2004-653294/63.  
 DR  
 XX Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for

19-APR-2002; 2002US-0373815P.  
 19-APR-2002; 2002US-0373817P.  
 19-APR-2002; 2002US-0373826P.  
 19-APR-2002; 2002US-0373884P.  
 22-APR-2002; 2002US-0374977P.  
 16-MAY-2002; 2002US-0381037P.  
 16-MAY-2002; 2002US-0381038P.  
 16-MAY-2002; 2002US-0381042P.  
 17-MAY-2002; 2002US-0381642P.  
 28-MAY-2002; 2002US-0383656P.  
 29-MAY-2002; 2002US-0383831P.  
 25-JUN-2002; 2002US-0391335P.  
 01-OCT-2002; 2002US-00262511.  
 (CURA-) CURAGEN CORP.  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Patturajan M, Spytke KA, Edinger SR, Ellerman K, Malyankar UM;  
 Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
 Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 WPI; 2003-381626/36.  
 DR N-PSDB; ADA05739.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 171; 586pp; English.  
 XX  
 CC The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
 CC and antilipaeic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.  
 XX  
 SQ Sequence 97 AA;

Query Match 100.0%; Score 49; DB 6; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 , Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMEYTVHL 9  
 Db 53 KMEYTVHL 61  
 RESULT 7  
 ADN62904  
 ID ADN62904 standard; protein; 97 AA.  
 XX  
 AC ADN62904;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DB Human NOV18e.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 29-OCT-2001; 2001US-0343629P.  
 PR 01-NOV-2001; 2001US-0349575P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 22-APR-2002; 2002US-0373884P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 28-MAY-2002; 2002US-0381642P.  
 PR 29-MAY-2002; 2002US-0383656P.  
 PR 25-JUN-2002; 2002US-0383831P.  
 XX  
 (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.

(GORM/) GORMAN L.  
 PA (ZERR/) ZERRHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG W.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIPI/) DIPIPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.  
 XX  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerrhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 WPI; 2004-213931/20.  
 DR N-PSDB; ADN62903.  
 DR  
 XX  
 Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 PT  
 XX  
 Claim 1; SEQ ID NO 100; 395pp; English.  
 PS  
 XX  
 The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX  
 SQ Sequence 97 AA;  
 Query Match 100.0%; Score 49; DB 8; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNEYTVHL 9

Db 53 KNEYTVHL 61  
 |||||  
 RESULT 8  
 ABG23378  
 ID ABG23378 standard; protein; 136 AA.  
 XX  
 AC ABG23378;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #23369.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS87565.  
 DR  
 XX  
 New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PT  
 PS Claim 20; SEQ ID NO 53737; 103pp; English.  
 XX  
 The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 136 AA;

Query Match 100.0%; Score 49; DB 4; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KNEYTVHL 9  
 |||||  
 Db 78 KNEYTVHL 86

```

KW chymotryptic enzyme; scce; enzyme.
XX Homo sapiens.
XX
XX US2003199010-A1.
XX
XX 23-OCT-2003.
XX
XX 13-JUN-2003; 2003US-00461787.
XX
XX 21-AUG-1997; 97US-00915659.
XX 21-AUG-1998; 98US-00137944.
XX 18-JUL-2000; 2000US-00618259.
XX
XX (UVAR-) UNIV ARKANSAS.
XX
XX O'brien TJ, Underwood LJ;
XX WPI; 2004-141550/14.
XX
XX Novel isolated DNA encoding tumor antigen derived gene-14 protein, useful
XX for treating neoplastic state (such as ovarian cancer, breast cancer,
XX lung cancer, colon cancer, prostate cancer) in an individual.
XX
XX Example 1; SEQ ID NO 4; 46pp; English.
XX
XX The invention relates to extracellular serine protease termed tumour
XX antigen derived gene-14 (TAGD-14) and its nucleic acid. Composition
XX comprising TAGD-14 peptide is useful for treating a neoplastic state in
XX an individual. The neoplastic state is chosen from ovarian cancer, breast
XX cancer, lung cancer, colon cancer, prostate cancer in which TAGD-14 is
XX overexpressed. The present sequence is Stratum corneum chymotryptic
XX enzyme (scce) catalytic domain. This sequence is used in the
XX exemplification of the invention.
XX
XX Sequence 144 AA;
SQ
Query Match 100.0%; Score 49; DB 8; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 9 KMNEYTVHL 17

RESULT 10
ADA05738
ID ADA05738 standard; protein; 181 AA.
XX
XX ADA05738;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human NOVI8d protein SEQ ID NO:98.
XX
XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
XX immunomodulator; cytostatic; nootropic; neuroprotective;
XX antiparkinsonian; antilipemic; gene therapy; human disease;
XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX immune disorder; haematopoietic disorder; dyslipidaemia.
XX
XX Homo sapiens.
XX
XX WO2003029424-A2.
XX
XX 10-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031373.
XX
XX 02-OCT-2001; 2001US-0326483P.
XX 05-OCT-2001; 2001US-0327435P.

```

```

KW chymotryptic enzyme; scce; enzyme.
XX Homo sapiens.
XX
XX US2003199010-A1.
XX
XX 23-OCT-2003.
XX
XX 13-JUN-2003; 2003US-00461787.
XX
XX 21-AUG-1997; 97US-00915659.
XX 21-AUG-1998; 98US-00137944.
XX 18-JUL-2000; 2000US-00618259.
XX
XX (UVAR-) UNIV ARKANSAS.
XX
XX O'brien TJ, Underwood LJ;
XX WPI; 2004-141550/14.
XX
XX Novel isolated DNA encoding tumor antigen derived gene-14 protein, useful
XX for treating neoplastic state (such as ovarian cancer, breast cancer,
XX lung cancer, colon cancer, prostate cancer) in an individual.
XX
XX Example 1; SEQ ID NO 4; 46pp; English.
XX
XX The invention relates to extracellular serine protease termed tumour
XX antigen derived gene-14 (TAGD-14) and its nucleic acid. Composition
XX comprising TAGD-14 peptide is useful for treating a neoplastic state in
XX an individual. The neoplastic state is chosen from ovarian cancer, breast
XX cancer, lung cancer, colon cancer, prostate cancer in which TAGD-14 is
XX overexpressed. The present sequence is Stratum corneum chymotryptic
XX enzyme (scce) catalytic domain. This sequence is used in the
XX exemplification of the invention.
XX
XX Sequence 144 AA;
SQ
Query Match 100.0%; Score 49; DB 8; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 9 KMNEYTVHL 17

RESULT 10
ADA05738
ID ADA05738 standard; protein; 181 AA.
XX
XX ADA05738;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human NOVI8d protein SEQ ID NO:98.
XX
XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
XX immunomodulator; cytostatic; nootropic; neuroprotective;
XX antiparkinsonian; antilipemic; gene therapy; human disease;
XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX immune disorder; haematopoietic disorder; dyslipidaemia.
XX
XX Homo sapiens.
XX
XX WO2003029424-A2.
XX
XX 10-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031373.
XX
XX 02-OCT-2001; 2001US-0326483P.
XX 05-OCT-2001; 2001US-0327435P.

```

05-OCT-2001; 2001US-0327449P.  
 09-OCT-2001; 2001US-0327417P.  
 09-OCT-2001; 2001US-0328029P.  
 09-OCT-2001; 2001US-0328044P.  
 09-OCT-2001; 2001US-0328056P.  
 12-OCT-2001; 2001US-0328849P.  
 15-OCT-2001; 2001US-0329414P.  
 17-OCT-2001; 2001US-0330142P.  
 18-OCT-2001; 2001US-0330309P.  
 22-OCT-2001; 2001US-0341058P.  
 24-OCT-2001; 2001US-0339266P.  
 24-OCT-2001; 2001US-0343629P.  
 29-OCT-2001; 2001US-0349575P.  
 01-NOV-2001; 2001US-0346357P.  
 17-APR-2002; 2002US-0373260P.  
 19-APR-2002; 2002US-0373815P.  
 19-APR-2002; 2002US-0373817P.  
 19-APR-2002; 2002US-0373884P.  
 22-APR-2002; 2002US-0374977P.  
 16-MAY-2002; 2002US-0381037P.  
 16-MAY-2002; 2002US-0381038P.  
 17-MAY-2002; 2002US-0381642P.  
 28-MAY-2002; 2002US-0383656P.  
 29-MAY-2002; 2002US-0383831P.  
 25-JUN-2002; 2002US-0391335P.  
 01-OCT-2002; 2002US-00262511.  
 (CURA-) CURAGEN CORP.  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Patcurajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 Ott T, Gorman L, Zernusen BD, Anderson DM, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiPippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 WPI; 2003-381626/36.  
 N-PSDB; ADA05737.  
 New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 pharmacogenomics.  
 Claim 1; Page 171; 586pp; English.  
 The present invention describes NOVX proteins, where X can be 1 to 55  
 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 described above and a carrier; (2) a kit comprising, in one or more  
 containers, the composition described above; (3) an isolated nucleic acid  
 molecule which encodes a NOVX protein of the invention; (4) a vector  
 comprising the nucleic acid molecule described above; (5) a cell  
 comprising the above vector; (6) an antibody that immunospecifically  
 binds to the polypeptide described above; (7) methods for determining the  
 presence or amount of the above polypeptide or nucleic acid molecule in a  
 sample; (8) methods for determining the presence of or predisposition to  
 a disease associated with altered levels of expression of the above  
 polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 method of identifying an agent that binds to the polypeptide described  
 above; (10) a method for identifying a potential therapeutic agent for  
 use in treating a pathology that is related to an aberrant expression or  
 aberrant physiological interactions of the polypeptide; (11) a method of  
 screening for a modulator of activity or of latency or predisposition to  
 a pathology associated with the polypeptide; (12) a method for modulating  
 the activity of the polypeptide described above; (13) methods of treating  
 or preventing a pathology associated with the above polypeptide in a  
 mammal; and (14) a method for producing the above polypeptide. NOVX  
 sequences have antidiabetic, anorectic, antibacterial, virucide,  
 immunomodulator, cytosstatic, nootropic, neuroprotective, antiparkinsonian  
 and antilipase activities, and can be used in gene therapy. The  
 polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.  
 XX  
 SQ Sequence 181 AA;  
 Query Match 100.0%; Score 49; DB 6; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KQNEYTVHL 9  
 Db 53 KQNEYTVHL 61  
 RESULT 12  
 ADN62902  
 ID ADN62902 standard; protein; 181 AA.  
 XX  
 AC ADN62902;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human NOV18d.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 12-OCT-2001; 2001US-0328056P.  
 PR 15-OCT-2001; 2001US-032849P.  
 PR 17-OCT-2001; 2001US-0329414P.  
 PR 18-OCT-2001; 2001US-0330142P.  
 PR 22-OCT-2001; 2001US-0330309P.  
 PR 24-OCT-2001; 2001US-0331058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 29-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.

25-JUN-2002; 2002US-0391335P.

(SMIT/) SMITHSON G.  
 (MILL/) MILLET I.  
 (PEYM/) PEYMAN J A.  
 (KEKU/) KEKUDA R.  
 (JUJ/) JU J.  
 (LILL/) LI L.  
 (GUOX/) GUO X.  
 (PATT/) PATTURAJAN M.  
 (SPYT/) SPYTEK K A.  
 (EDIN/) EDINGER S R.  
 (ELLE/) ELLERMAN K.  
 (MALY/) MALYANKAR U M.  
 (ORTT/) ORT T.  
 (GORM/) GORMAN L.  
 (ZERH/) ZERHUSEN B D.  
 (ANDE/) ANDERSON D W.  
 (ZHON/) ZHONG M.  
 (CATT/) CATTERTON E.  
 (JIWW/) JI W.  
 (MILL/) MILLER C E.  
 (RAST/) RASTELLI L.  
 (STON/) STONE D J.  
 (PENA/) PENNA C E A.  
 (SHEN/) SHENOY S G.  
 (SHIM/) SHIMKETS R A.  
 (ROTH/) ROTHENBERG M E.  
 (LEAC/) LEACH M D.  
 (AGEE/) AGEE M L.  
 (BERG/) BERGHS C.  
 (DIPI/) DIPIPO V A.  
 (EISE/) EISEN A.  
 (GANG/) GANGOLLI E A.  
 (RIEG/) RIEGER D K.  
 (SPAD/) SPADERNA S K.

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X; Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM; Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E; Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG; Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA; Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

WPI; 2004-213931/20.  
 N-PSDB; ADN62301.

Isolated NOVX polypeptides and nucleic acids, useful for preventing, diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

Claim 1; SEQ ID NO 98; 395pp; English.

The invention relates to isolated NOVX polypeptides and polynucleotides. NOVX polypeptides and polynucleotides are used to prevent, diagnose or treat a medical condition in human related to the aberrant expression and activity of NOVX polypeptides. For example, NOVX polypeptides and polynucleotides may be used to treat disorders associated with decreased expression or activity of NOVX by supplementing the patient our production or to rectify mutations. Conversely, antisense NA molecules may be administered to down regulate expression of NOVX polypeptides by binding with the cells own genes and preventing their expression. NOVX polynucleotides and complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar sequences in samples, and so which patients may be in need of restorative therapy. NOVX polypeptides may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of NOVX. The anti-NOVX polypeptide antibodies, agonists and antagonists may also be used to modulate NOVX polynucleotide expression and activity of NOVX polypeptides. The anti-NOVX polypeptide antibodies may also be used as diagnostic agents for detecting the presence of NOVX in samples. NOVX polypeptides and polynucleotides may be used in this way to prevent, diagnose and treat: metabolic disorders, diabetes, obesity, infectious

CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers. CC They may also be used as antibacterial agents. The present sequence CC represents the amino acid sequence of a human NOVX protein.

XX  
 SQ Sequence 181 AA;

Query Match 100.0%; Score 49; DB 8; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 0.22; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

QY 1 KMNEYTVHL 9  
 DB 53 KMNEYTVHL 61  
 |||||

RESULT 13  
 ADA05736

ID ADA05736 standard; protein; 198 AA.

XX  
 AC ADA05736;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV18c protein SEQ ID NO:96.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX  
 OS Homo sapiens.

XX  
 PN WO2003029424-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 02-OCT-2002; 2002WO-US031373.

XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Smithson G., Millet I., Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rotherberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 DR WPI; 2003-381626/36.  
 DR N-PSDB; ADA05735.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 170; 586pp; English.  
 XX  
 CC The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOVX). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
 CC and antipsychotic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.  
 XX  
 SQ Sequence 198 AA;  
 Query Match 100.0%; Score 49; DB 6; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KQNEYTVHL 9  
 Db 72 KQNEYTVHL 80  
 RESULT 14  
 ADN62900  
 ID ADN62900 standard; protein; 198 AA.  
 XX  
 AC ADN62900;  
 XX

DT 01-JUL-2004 (first entry)  
 XX Human NOV18c.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327445P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENNA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.  
PA (LEAC/) LEACH M D.  
PA (AGEE/) AGE M L.  
PA (BERG/) BERGHS C.  
PA (DIP/) DIPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerkhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
XX WPI; 2004-213931/20.  
DR N-PSDB; ADN62899.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
PS Claim 1; SEQ ID NO 96; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
XX represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 198 AA;  
Query Match 100.0%; Score 49; DB 8; Length 198;  
Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Gaps 0;  
Matches 9; Conservative 0; Indels 0;

OY 1 KKNKYTVHL 9  
Db 72 KKNKYTVHL 80  
|||||  
|||||

RESULT 15  
ADA05744  
ID ADA05744 standard; protein; 224 AA.  
XX  
AC ADA05744;  
XX  
XX 06-NOV-2003 (first entry)  
DT  
XX

DE Human NOV18g protein SEQ ID NO:104.  
XX  
KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
XX 02-OCT-2002; 2002WO-US031373.  
XX  
XX 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
XX Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
XX Ort T, Gorman L, Zerkhusen BD, Anderson DW, Zhong M, Catterton E;  
XX Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
XX Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
XX Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
XX WPI; 2003-381626/36.  
DR N-PSDB; ADA05743.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Claim 1; Page 172; 586pp; English.  
XX  
XX The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell



comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity or of latency or predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide. NOVX sequences have antidiabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian and antilipidemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a syndrome associated with a human disease. The polypeptide or the nucleic acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence represents a human NOVX from the present invention.

XX SQ Sequence 224 AA;

Query Match 100.0%; Score 49; DB 6; Length 224;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KXNEYTVHL 9  
Db 58 KXNEYTVHL 66  
|||||||

Search completed: July 13, 2005, 17:19:49  
Job time : 78.6667 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083A-31

Perfect score: 49

Sequence: 1 KMNEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	253	2	serine proteinase
2	39	79.6	218	1	lysis protein t -
3	39	79.6	218	1	lysis protein t -
4	37	75.5	334	2	conserved hypothet
5	37	75.5	393	2	hypothetical prote
6	36	73.5	627	2	hypothetical prote
7	36	73.5	1829	2	hypothetical prote
8	35	71.4	159	2	spermidine/spermin
9	35	71.4	159	2	spermidine/spermin
10	35	71.4	505	2	conserved hypothet
11	35	71.4	555	2	acylase and dieste
12	35	71.4	667	2	ring finger protei
13	35	71.4	667	2	ring finger protei
14	34	69.4	270	2	replication protei
15	34	69.4	270	2	replication protei
16	34	69.4	285	2	hypothetical prote
17	34	69.4	421	2	B. subtilis twbN p
18	34	69.4	421	2	conserved hypothet
19	33	67.3	204	2	FixJ Transcription
20	33	67.3	204	2	nitrogen fixation
21	33	67.3	218	2	calflagin Tb-24 -
22	33	67.3	218	2	hypothetical prote
23	33	67.3	227	2	two-component resp
24	33	67.3	229	2	calflagin Tb-1.7 -
25	33	67.3	233	1	flagellar calcium -
26	33	67.3	407	2	calflagin Tb-44A -
27	33	67.3	422	2	hypothetical prote
28	33	67.3	429	2	preprotein translo
29	33	67.3	525	2	oligopeptide ABC t

30	33	67.3	664	2	C72379
31	33	67.3	677	2	A87470
32	33	67.3	1042	2	S23738
33	33	67.3	1308	2	E71622
34	33	67.3	1481	1	OZDOP3
35	32	65.3	49	2	H89908
36	32	65.3	84	2	A82172
37	32	65.3	116	2	S24989
38	32	65.3	130	2	E37262
39	32	65.3	144	2	F71215
40	32	65.3	192	2	G81301
41	32	65.3	228	2	C90033
42	32	65.3	304	2	H75378
43	32	65.3	339	2	F97121
44	32	65.3	371	2	A89800
45	32	65.3	448	2	F82280

#### ALIGNMENTS

##### RESULT 1

A53968

serine proteinase SCCE precursor - human

N:Alternate names: stratum corneum chymotryptic enzyme

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A53968

R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A:Reference number: A53968; MUID:94308225; PMID:8034709

A:Accession: A53968

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <HAN>

A:Cross-references: UNIPROT:P49862; GB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G532504

C:Genetics:

A:Gene: GDB:PRSS6; SCCE

A:Cross-references: GDB:377730

A:Map position: 7q35-7q35

C:Superfamily: trypsin; trypsin homology

F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 49; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

DB 72 KMNEYTVHL 80

##### RESULT 2

YVBP3

lysis protein t - phase K3

C:Species: phase K3

A:Note: host Escherichia coli

C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004

C:Accession: A27083

R:Riede, I.

J. Bacteriol. 169, 2956-2961, 1987

A:Title: Lysis Gene t of T-even bacteriophages: evidence that colicins and bacteriophage

A:Reference number: A27083; MUID:87250254; PMID:3597316

A:Accession: A27083

A:Molecule type: DNA

A:Residues: 1-218 <RIB>

A:Cross-references: UNIPROT:P10393; GB:M16812; NID:G515503; PIDN:AAA88415.1; PID:G215504

A:Note: The author translated the codon CAA for residue 85 as Ile and CAG for residue 209;

C:Genetics:

A:Gene: t

C:Superfamily: phase T4 lysis protein t

C:Keywords: host cell lysis

```
Query Match      79.6%; Score 39; DB 1; Length 218;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 MNEYTVHL 9
      |:|||||
Db      139 MDEYTVHL 146

RESULT 3
YBPT4
Lysis protein t - phage T4
C:Species: phage T4
A:Note: host Escherichia coli
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: JF0028; S07395
R:Montag, D.; Degen, M.; Henning, U.
Nucleic Acids Res. 15, 6736, 1987
A:Title: Nucleotide sequence of gene t (lysis gene) of the E. coli phage T4.
A:Reference number: S07395; MUID:87316934; PMID:3628006
A:Accession: JF0028
A:Molecule type: DNA
A:Residues: 1-218 <MON>
A:Cross-references: UNIPROT:P06808; GB:Y00408; NID:G15368; PIDN:CAA68470.1; PID:G15369
A:Note: the sequence is almost identical with that of the E.coli phage K3
C:Comment: At the end of the growth cycle, phage T4 expresses two genes with lysis function about the gene product of t, although it has been suggested that it acts as a phospholipase.
C:Genetics:
A:Gene: t
A:Map position: 157.985-158.639
C:Superfamily: phage T4 lysis protein t
C:Keywords: host cell lysis; transmembrane protein
F:35-49/Domain: transmembrane #status predicted <TMN>

Query Match      79.6%; Score 39; DB 1; Length 218;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 MNEYTVHL 9
      |:|||||
Db      139 MDEYTVHL 146

RESULT 4
B98019
conserved hypothetical protein sprl179 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98019
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Ee, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
Y. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; et al.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B98019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: UNIPROT:Q8DPG2; GB:AE007317; PIDN:AAK99982.1; PID:G15458811; GSPDB:G15458811
C:Genetics:
A:Gene: sprl179

Query Match      75.5%; Score 37; DB 2; Length 334;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 KMNEYTVHL 9
      |:|||||
Db      302 KMNEYTVHL 310
```

```
RESULT 5
D75207
hypothetical protein PAB2235 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: D75207
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: D75207
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <KAW>
A:Cross-references: UNIPROT:Q9V291; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CAB4910;
A:Experimental source: strain Orsay
C:Genetics:
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0196

Query Match      75.5%; Score 37; DB 2; Length 393;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 KMNEYTVHL 9
      |:|||||
Db      129 KLPEYTVHL 137

RESULT 6
T25395
hypothetical protein T28A8.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25395
R:Lloyd, C.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z20027
A:Accession: T25395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-627 <WIL>
A:Cross-references: UNIPROT:Q9XU04; EMBL:Z92813; PIDN:CAB07289.1; GSPDB:GN00021; CESP:T25395
A:Experimental source: clone T28A8
C:Genetics:
A:Gene: CESP:T28A8.6
A:Map position: 3
A:Introns: 51/2; 89/2; 183/2; 221/2; 296/2; 607/3
C:Superfamily: Caenorhabditis elegans hypothetical protein Y57A10A.d

Query Match      73.5%; Score 36; DB 2; Length 627;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 KMNEYTVHL 9
      |:|||||
Db      62 KMNEYSVIEL 70

RESULT 7
T34239
hypothetical protein F26F12.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34239
R:Wilson, R.; Bentley, D.; Gattung, S.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F26F12.
A:Reference number: Z21493
A:Accession: T34239
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
```

A;Residues: 1-1829 <WIL>  
A;Cross-references: UNIPROT:Q19815; EMBL:U55373; PIDN:AC25894.1; GSPDB:GN00023; CESP:F2  
A;Experimental source: strain Bristol NZ; clone F26F12  
C;Genetics:  
A;Gene: CESP:F26F12.7  
A;Map position: 5  
A;Introns: 110/3; 441/3; 801/2; 1244/3; 1693/2; 1784/1

Query Match 73.5%; Score 36; DB 2; Length 1829;  
Best Local Similarity 71.4%; Pred. No. 97;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KMNEYTVH 8  
|||:|  
Db 765 LNEYTTH 771

RESULT 8  
A11782  
spermidine/spermine N1-acetyl transferase homolog lmo2658 [imported] - Listeria monocytogenes  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: A11406  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: A11406  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-159 <GLA>  
A;Cross-references: UNIPROT:Q8Y419; GB:NC\_003210; PIDN:CAD00871.1; PID:g16412158; GSPDB:G  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo2658

Query Match 71.4%; Score 35; DB 2; Length 159;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVH 8  
|||:|  
Db 139 KMSEWTVH 146

RESULT 9  
A11782  
spermidine/spermine N1-acetyl transferase homolog lin2807 [imported] - Listeria innocua  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: A11782  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: A11782  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-159 <GLA>  
A;Cross-references: UNIPROT:Q92711; GB:AL592022; PIDN:CAC98033.1; PID:g16415343; GSPDB:G  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin2807

Query Match 71.4%; Score 35; DB 2; Length 159;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVH 8  
|||:|  
Db 139 KMSEWTVH 146

RESULT 10  
H75431  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: H75431  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F.M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; May, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: AY5250; MUID:20036896; PMID:10567286  
A;Accession: H75431  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-505 <WHI>  
A;Cross-references: UNIPROT:Q9RV79; GB:AE001964; GB:AE000513; NID:g6458881; PIDN:AAF1072;  
A;Experimental source: strain R1  
C;Genetics:  
A;Gene: DR1150  
A;Map position: 1  
C;Superfamily: conserved hypothetical protein b0835

Query Match 71.4%; Score 35; DB 2; Length 505;  
Best Local Similarity 66.7%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9  
:|||||  
Db 57 QMNEYDTHL 65

RESULT 11  
AD1794  
acylase and diesterase homolog lin2898 [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AD1794  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AD1794  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-555 <GLA>  
A;Cross-references: UNIPROT:Q926Z1; GB:AL592022; PIDN:CAC98124.1; PID:g16415433; GSPDB:G  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin2898

Query Match 71.4%; Score 35; DB 2; Length 555;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9  
:|||||  
Db 483 KINEYTDL 491

RESULT 12

T09013  
RING finger protein Fxy - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 04-Apr-2004  
C:Accession: T09013  
R:Palmer, S.; Perry, J.; Kipling, D.; Ashworth, A.  
Proc. Natl. Acad. Sci. U.S.A. 94, 12030-12035, 1997  
A:Title: A gene spans the pseudautosomal boundary in mice.  
A:Reference number: Z16531; MUID:98004518; PMID:9342357  
A:Accession: T09013  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-667 <PAL>  
A:Cross-references: EMBL:AF026565; NID:g2589222; PIDN:AAB83986.1; PID:g2589223  
C:Genetics:  
A:Gene: Fxy  
A:Map position: X; Y  
C:Superfamily: rfp transforming protein  
C:Keywords: zinc finger  
F:6-65/Domain: RING finger homology <RRN>  
  
Query Match 71.4%; Score 35; DB 2; Length 667;  
Best Local Similarity 75.0%; Pred. No. 53;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 KMNEYTVH 8  
Db 444 KQHVTVH 451  
  
RESULT 13  
T09482  
RING finger protein FXY - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T09482  
R:Perry, J.; Feather, S.; Smith, A.; Palmer, S.; Ashworth, A.  
submitted to the EMBL Data Library, November 1997  
A:Description: The human FXY maps to chromosome Xp22.3: Implications for evolution of the  
A:Reference number: Z16687  
A:Accession: T09482  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-667 <PER>  
A:Cross-references: UNIPROT:O15344; EMBL:AF035360; NID:g2827993; PID:g2827994  
C:Genetics:  
A:Gene: FXY  
A:Map position: Xp22.3  
C:Superfamily: rfp transforming protein  
F:6-65/Domain: RING finger homology <RRN>  
  
Query Match 71.4%; Score 35; DB 2; Length 667;  
Best Local Similarity 75.0%; Pred. No. 53;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 KMNEYTVH 8  
Db 444 KQHVTVH 451  
  
RESULT 14  
A43711  
replication protein repA 32K chain - human  
C:Species: Homo sapiens (man)  
C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 09-Jul-2004  
C:Accession: A43711  
R:Erdile, L.F.; Wold, M.S.; Kelly, T.J.  
J. Biol. Chem. 265, 3177-3182, 1990  
A:Title: The primary structure of the 32-kDa subunit of human replication protein A.  
A:Reference number: A43711; MUID:90153966; PMID:2406247  
A:Accession: A43711  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-270 <ERD>  
A:Cross-references: UNIPROT:P15927; EMBL:J05249; NID:g337349; PIDN:AAA36560.1; PID:g33733  
C:Superfamily: Schizosaccharomyces pombe single-stranded DNA binding protein 30K chain  
  
Query Match 69.4%; Score 34; DB 2; Length 270;  
Best Local Similarity 62.5%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 MNEYTVHL 9  
Db 152 MNEFTTHI 159  
  
RESULT 15  
S28682  
replication protein A 32K chain homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C:Accession: S28682  
R:Nakagawa, M.; Tsukada, S.; Soma, T.; Shimizu, Y.; Miyake, S.; Iwamatsu, A.; Sugiyama, T.  
Nucleic Acids Res. 19, 4292, 1991  
A:Title: cDNA cloning of the murine 30-kDa protein homologous to the 32-kDa subunit of h  
A:Reference number: S28682; MUID:91334146; PMID:1908076  
A:Accession: S28682  
A:Molecule type: mRNA  
A:Residues: 1-270 <NAG>  
A:Cross-references: UNIPROT:Q62193; EMBL:D00812; NID:g220583; PIDN:BAA00693.1; PID:g22058  
C:Superfamily: Schizosaccharomyces pombe single-stranded DNA binding protein 30K chain  
  
Query Match 69.4%; Score 34; DB 2; Length 270;  
Best Local Similarity 62.5%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 MNEYTVHL 9  
Db 152 MNEFTTHI 159  
  
Search completed: July 13, 2005, 17:31:24  
Job time : 15.5556 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-31

Perfect score: 49

Sequence: 1 KMNEVTVHL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_Q3.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	253	1 KLK7_HUMAN	P49862 homo sapien
2	49	100.0	253	2 Q8NSN9	Q8NSN9 homo sapien
3	44	89.8	181	2 Q8NFV7	Q8NFV7 homo sapien
4	39	79.6	210	2 Q9G0B3	Q9G0B3 bacterioph
5	39	79.6	218	1 VLYS_BPK3	P10393 bacterioph
6	39	79.6	218	1 VLYS_BPT4	P06808 bacterioph
7	38	77.6	261	1 RFA4_HUMAN	Q13156 homo sapien
8	37	75.5	334	2 Q8DPG2	Q8DPG2 streptococc
9	37	75.5	393	1 PRIL_PPRAB	Q9V291 pyrococcus
10	37	75.5	936	2 Q633P6	Q633P6 bacillus ce
11	36	73.5	445	2 Q67218	Q67218 arabidopsis
12	36	73.5	603	2 Q8L836	Q8L836 arabidopsis
13	36	73.5	627	2 Q9XU04	Q9XU04 caenorhabdi
14	36	73.5	635	2 Q9LQW1	Q9LQW1 arabidopsis
15	36	73.5	753	2 Q8VDV0	Q8VDV0 streptococc
16	36	73.5	753	2 Q8E174	Q8E174 streptococc
17	36	73.5	753	2 Q8E6M8	Q8E6M8 caenorhabdi
18	36	73.5	1829	2 Q19815	Q19815 caenorhabdi
19	35	71.4	47	2 Q80274	Q80274 mus macedon
20	35	71.4	47	2 Q80277	Q80277 mus cervico
21	35	71.4	47	2 Q80280	Q80280 mastomys hu
22	35	71.4	47	2 Q80283	Q80283 arvicola te
23	35	71.4	47	2 Q802B4	Q802B4 mus platyth
24	35	71.4	47	2 Q810Y2	Q810Y2 mus spicile
25	35	71.4	48	2 Q810X9	Q810X9 mus spretus
26	35	71.4	159	2 Q8V419	Q8V419 listeria mo
27	35	71.4	159	2 Q92711	Q92711 listeria in
28	35	71.4	159	2 Q71WB4	Q71WB4 listeria mo
29	35	71.4	197	2 Q775S2	Q775S2 cryptophleb
30	35	71.4	400	2 Q6PD02	Q6PD02 mus musculu
31	35	71.4	505	2 Q9RV79	Q9RV79 deinococcus

32	35	71.4	555	2	Q92621	Q92621 listeria in
33	35	71.4	579	2	Q8IBMS	Q8IBMS plasmodium
34	35	71.4	667	1	MID1_HUMAN	O15344 homo sapien
35	35	71.4	667	1	MID1_MUSSP	P82457 mus spretus
36	35	71.4	667	1	MID1_RAT	P82458 rattus norv
37	35	71.4	667	2	Q90WD1	Q90WD1 gallus gall
38	35	71.4	667	2	Q71R46	Q71R46 gallus gall
39	35	71.4	668	2	Q6DEU6	Q6DEU6 xenopus tro
40	35	71.4	680	1	MID1_MOUSE	Q70583 mus musculu
41	35	71.4	685	1	MID2_MOUSE	Q9QU86 mus musculu
42	35	71.4	685	2	Q6GX19	Q6GX19 aotus trivi
43	35	71.4	685	2	Q6GX20	Q6GX20 cercopithe
44	35	71.4	687	2	Q8IE27	Q8IE27 plasmodium
45	35	71.4	715	1	MID2_HUMAN	Q9JUV3 homo sapien

#### ALIGNMENTS

RESULT 1  
 ID KLK7\_HUMAN STANDARD; PRT; 253 AA.  
 AC P49862;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).  
 GN Name=KLK7; Synonyms=PRSS6, SCCE;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
 RC TISSUE=Skin;  
 RX MEDLINE=94308225; PubMed=8034709;  
 RA Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;  
 RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";  
 RL J. Biol. Chem. 269:19420-19426(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=keratinocytes;  
 RA Yousef G.M., Scorilas A., Diamandis E.P.;  
 RT "Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepker B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J., Wallbrandt P., Egelrud T.;  
 RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in mice; a model for chronic itchy dermatitis.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95314630; PubMed=7794273;  
 RA Skjott A., Stroemqvist M., Egelrud T.;  
 RT "Primary substrate specificity of recombinant human stratum corneum chymotryptic enzyme.";  
 RL Biochem. Biophys. Res. Commun. 211:586-589(1995).  
 CC !- FUNCTION: May catalyze the degradation of intercellular cohesive structures in the cornified layer of the skin in the continuous

sheddung of cells from the skin surface. Specific for amino acid residues with aromatic side chains in the P1 position. SCCE cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-Phe-|-Tyr-26 and 26-Tyr-|-Thr-27. Could play a role in the activation of precursors to inflammatory cytokines.

-!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is expressed by keratinocytes in the epidermis. Very low levels are also seen in the brain and kidney.

-!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein subfamily.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; L33404; AAC37551.1; --  
DR EMBL; AF166330; AAD49718.1; --  
DR EMBL; AF243327; AAG33360.1; --  
DR EMBL; AF332583; AAG69624.1; --  
DR PIR; A53968; A53968.  
DR HSSP; P00760; 1EZX.  
DR MEROPS; S01.300; --  
DR Genew; HGNC:6368; KLK7.  
DR H-InvDB; HIX0015373; --  
DR MIM; 604438; --  
DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
DR GO; GO:0008544; P:epidermal differentiation; TAS.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
KW Signal; Zymogen.

FT SIGNAL 1 22 Activation peptide.  
FT PROPEP 23 29 Kallikrein 7.  
FT CHAIN 30 253  
FT ACT\_SITE 70 70 Charge relay system (By similarity).  
FT ACT\_SITE 112 112 Charge relay system (By similarity).  
FT ACT\_SITE 205 205 Charge relay system (By similarity).  
FT DISULFID 36 137 By similarity.  
FT DISULFID 55 71 By similarity.  
FT DISULFID ? 239 By similarity.  
FT DISULFID 144 211 By similarity.  
FT DISULFID 176 190 By similarity.  
FT DISULFID 201 226 By similarity.  
FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;

Query Match 100.0%; Score 49; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNYTVHL 9  
Db 72 KKNYTVHL 80  
|||||

RESULT 2  
QBN5N9 PRELIMINARY; PRT; 253 AA.  
AC QBN5N9  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE GN Name=KLK7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.M., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; BC032005; AAH32005.1; --  
DR HSSP; P00760; 1EZX.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 100.0%; Score 49; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNYTVHL 9  
Db 72 KKNYTVHL 80  
|||||

RESULT 3  
QBNFV7 PRELIMINARY; PRT; 181 AA.  
AC QBNFV7  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Kallikrein 7 short variant protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovarian carcinoma;
RX MEDLINE=22623266; PubMed=12738725;
RA Dong Y., Kaushal A., Brattstrand M., Nicklin J., Clements J.A.;
RT "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer
RT produces novel variants with potential as cancer biomarkers.";
RL Clin. Cancer Res. 9:1710-1720(2003).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF411215; AAN03663.1; -.
DR HSSP; P00760; 1EXZ.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 181 AA; 19887 MW; 86A28A03880C2D78 CRC64;

Query Match 89.8%; Score 44; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MNEYTVHL 9
Db 1 MNEYTVHL 8

RESULT 4
Q9G0B3 PRELIMINARY; PRT; 210 AA.
AC Q9G0B3
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE T protein (Fragment).
OS Bacteriophage AR1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=66711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20485545; PubMed=11029414;
RX DOI=10.1128/JB.182.21.5962-5968.2000;
RA Yu S.L., Ko K.L., Chen C.S., Chang Y.C., Syu W.J.;
RT "Characterization of the distal tail fiber locus and determination of
RT the receptor for phage AR1, which specifically infects Escherichia
RT coli O157:H7.";
RL J. Bacteriol. 182:5962-5968(2000).
DR EMBL; AF208841; AAG29756.1; -.
FT NON TER 210
FT SEQUENCE 210 AA; 24348 MW; C6FF585F9FE0DF68 CRC64;

Query Match 79.6%; Score 39; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MNEYTVHL 9
Db 139 MDEYTVHL 146

RESULT 5
VLVS_BPK3
AC P10393; STANDARD; PRT; 218 AA.
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lysis protein.
GN Name=L;
OS Bacteriophage K3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=10674;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87250254; PubMed=3597316;
RA Riede I.;
RT "Lysis gene t of T-even bacteriophages: evidence that colicins and
RT bacteriophage genes have common ancestors.";
RL J. Bacteriol. 169:2956-2961(1987).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M16812; AAA88415.1; -.
DR PIR; A27083; YVBPk3.
KW Phage lysis protein.
SQ SEQUENCE 218 AA; 25222 MW; 21B4DC02ACA0ECF6 CRC64;

Query Match 79.6%; Score 39; DB 1; Length 218;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MNEYTVHL 9
Db 139 MDEYTVHL 146

RESULT 6
VLVS_BPT4
AC P06808; STANDARD; PRT; 218 AA.
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lysis protein (Holin) (Protein rv).
GN Name=L; Synonyms=RV;
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87316934; PubMed=3628006;
RA Montag D., Degen M., Henning U.;
RT "Nucleotide sequence of gene t (lysis gene) of the E. coli phage T4.";
RL Nucleic Acids Res. 15:6736-6736(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2514363; PubMed=12626685; DOI=10.1128/MMBR.67.1.86-156.2003;
RA Miller E.S., Kutter E., Mosig G., Arisaka F., Kunisawa T., Ruger W.;
RT "Bacteriophage T4 genome.";
RL Microbiol. Mol. Biol. Rev. 67:86-156(2003).
RN [3]
RP SEQUENCE OF 1-34 FROM N.A.
RX MEDLINE=88011316; PubMed=2958637;
RA Montag D., Riede I., Eschbach M.-L., Degen M., Henning U.;
RT "Receptor-recognizing proteins of T-even type bacteriophages. Constant
RT and hypervariable regions and an unusual case of evolution.";
RL J. Mol. Biol. 196:165-174(1987).
RN [4]

```

```
RP SEQUENCE OF 202-218 FROM N.A.
RC STRAIN=D;
RX MEDLINE=93106978; PubMed=8416914;
RA Orsini G., Ouhammouch M., le Caer J.P., Brody E.N.;
RT "The asia gene of bacteriophage T4 codes for the anti-sigma 70
RT protein.";
RL J. Bacteriol. 175:85-93(1993).
CC -!- FUNCTION: At the end of the growth cycle, phase T4 expresses two
CC genes with lysis function, e and t. Nothing is known about the
CC gene product of t, although it has been suggested that it acts as
CC a phospholipase.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00408; CAA68470.1; -.
DR EMBL; AF158101; AAD42661.1; -.
DR EMBL; X05677; CAA29164.1; -.
DR EMBL; M99441; AAA32481.1; -.
DR FIR; JF0028; YVBPT4.
KW Phage lysis protein; Transmembrane.
FT TRANSMEM 35
FT TRANSMEM 49 Potential.
SQ SEQUENCE 218 AA; 25175 MW; 9110BE111D772DF5 CRC64;

Query Match 79.6%; Score 39; DB 1; Length 218;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MNEYTVHL 9
Db 139 MDEYTVHL 146

RESULT 7
RFA4 HUMAN
ID RFA4 HUMAN STANDARD; PRT; 261 AA.
AC Q13156;
DT 01-NOV-1997 (Rel. 35; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 25-OCT-2004 (Rel. 45; Last annotation update)
DE Replication protein A 30 kDa subunit (RP-A) (RP-A) (Replication
DE factor-A protein 4).
GN Name=RPA4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95280910; PubMed=7760808;
RA Keshav K.F., Chen C., Dutta A.;
RT "Rpa4, a homolog of the 34-kilodalton subunit of the replication
RT protein A complex.";
RL Mol. Cell. Biol. 15:3119-3128(1995).
RN [2]
RP REVISION TO 10.
RA Keshav K.F.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT THR-33.
RA Rieder M.J., Armet T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA O'Neil M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RT "NIH-SNPs, environmental genome project, NIHES ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Absolutely required for simian virus 40 DNA replication
CC in vitro. It participates in a very early step in initiation. RP-A
CC is a single-stranded DNA-binding protein (By similarity).
```

```
CC -!- SUBUNIT: Heterotrimer of 70, 32/30, and 14 kDa chains. The DNA-
CC binding activity may reside exclusively on the 70 kDa subunit.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Preferentially expressed in placental and
CC colon mucosa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U24186; AAB08488.2; -.
DR EMBL; AF494047; AAM09569.1; -.
DR HSPF; P15927; IQUQ.
DR Reactome; Q13156; -.
DR GO; GO:0005662; C:DNA replication factor A complex; TAS.
DR GO; GO:0003697; P:single-stranded DNA binding; TAS.
DR GO; GO:0006270; P:DNA replication initiation; TAS.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004365; tRNA_anti.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01336; tRNA_anti; 1.
KW Alternative splicing; DNA replication; Nuclear protein; Polymorphism.
FT VARIANT 33 33 A -> T.
FT VARIANT 33 33 /FTID=VAR_019170.
SQ SEQUENCE 261 AA; 28868 MW; 6A925FAEBE21718 CRC64;

Query Match 77.6%; Score 38; DB 1; Length 261;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MNEYTVHL 9
Db 151 MNEFTVHL 158

RESULT 8
Q8DPG2 PRELIMINARY; PRT; 334 AA.
AC Q8DPG2;
DT 01-MAR-2003 (TrEMBLrel. 23; Created)
DT 01-MAR-2003 (TrEMBLrel. 23; Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24; Last annotation update)
DE Hypothetical protein sprl179.
GN OrderedLocNames=sprl179;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geinger C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skarud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008490; AAK99982.1; -.
DR PIR; B98019; B98019.
DR InterPro; IPR006314; Dyp_peroxidase.
DR Pfam; PF04261; Dyp_perox; 1.
DR TIGRFAMs; TIGR01413; Dyp_perox_fam; 1.
```

KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 334 AA; 38137 MW; E207F1BC267334E5 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 334;  
Best Local Similarity 66.7%; Pred. No. 51;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
| | | | | | | | | |  
DB 302 KMNEYTVHL 310

RESULT 9  
PRIL PYRAB STANDARD; PRT; 393 AA.  
AC Q9V291;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE DNA primase large subunit (EC 2.7.7.-) (DNA primase 46 kDa subunit) (p46).  
DE (p46).  
GN Name=prIB; OrderedLocNames=PYRAB01830; ORFNames=PAB2235;  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GES / Orsay;  
RX MEDLINE=22511545; PubMed=12622808;  
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O., Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C., van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;  
RT "An integrated analysis of the genome of the hyperthermophilic archaeon Pyrococcus abyssi."  
RL Mol. Microbiol. 47:1495-1512(2003).  
CC -!- FUNCTION: DNA primase is the polymerase that synthesizes small RNA primers for the Okazaki fragments on both template strands at replication forks during chromosomal DNA synthesis (By similarity).  
CC -!- SUBUNIT: Heterodimer of a small subunit and a large subunit. Both participate in formation of the active center, but the ATP-binding site is exclusively located on the small subunit (By similarity).  
CC -!- SIMILARITY: Belongs to the eukaryotic-type primase large subunit family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AJ248283; CAB49107.1; --  
DR FIC; D75207; D75207.  
DR HAWAP; MF 00701; --; 1.  
DR InterPro; IPR008918; 5\_3\_exo.C.  
DR InterPro; IPR007238; DNA primase lrg.  
DR Pfam; PF04104; DNA primase lrg.1.  
KW Complete proteome; DNA replication; DNA-directed RNA polymerase; Primosome; Transference.  
SQ SEQUENCE 393 AA; 45485 MW; 9448642FBB3FBF43 CRC64;

Query Match 75.5%; Score 37; DB 1; Length 393;  
Best Local Similarity 66.7%; Pred. No. 61;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
| | | | | | | | | |  
DB 129 KLPEYTVHL 137

RESULT 12

RESULT 10  
Q633P6 PRELIMINARY; PRT; 936 AA.  
ID Q633P6  
AC Q633P6;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=BIZK4292;  
OS Bacillus cereus ZK.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=288681;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ZK;  
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K., Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R., Richardson P., Rubin E., Tice H.;  
RT "Complete genome sequence of Bacillus cereus ZK."  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; CP000001; AAU15977.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 936 AA; 104539 MW; D05C843D8B02DF8A CRC64;

Query Match 75.5%; Score 37; DB 2; Length 936;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVH 8  
| | | | | | | | | |  
DB 749 KMHDYTVH 756

RESULT 11  
Q67Z18 PRELIMINARY; PRT; 445 AA.  
ID Q67Z18  
AC Q67Z18;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE MRNA, partial cds, clone: RAFL23-25-J01 (MRNA, partial cds, clone: RAFL23-09-L14) (Fragment).  
DE Arabidopsis thaliana (Mouse-ear cress).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A., Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y., Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K., Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J., Hayashizaki Y., Shinozaki K.;  
RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK176300; BAD44063.1; --  
DR EMBL; AK176215; BAD43978.1; --  
DR InterPro; IPR007145; MAP65\_ASE1.  
DR Pfam; PF03999; MAP65\_ASE1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 445 AA; 51244 MW; 00BD7C51A0D6EA45 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 445;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
| | | | | | | | | |  
DB 11 KLNEYQTHL 19

```
Q9L836
ID Q8L836 PRELIMINARY; PRT; 603 AA.
AC Q8L836
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein Atg14690.
GN Name=Atg14690;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayaishizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayaishizaki Y., Hsuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY120768; AAM53326.1; -.
DR EMBL; BT008373; AAP37732.1; -.
DR InterPro; IPR007145; MAP65_ASE1.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF03999; MAP65_ASE1; 1.
KW Hypothetical protein.
SQ SEQUENCE 603 AA; 69060 MW; 7C2B37A23CBB889D CRC64;

Query Match 73.5%; Score 36; DB 2; Length 603;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9
|:|:|:|
Db 169 KLINEYQTHL 177

RESULT 13
Q9XU04 PRELIMINARY; PRT; 627 AA.
AC Q9XU04;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein T28A8.6.
GN ORFNames=T28A8.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lloyd C.R.;

Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z92813; CAB07289.1; -.
DR PIR; T25395; T25395
DR WormBase; WBGene00012112; T28A8.6.
DR WormPep; T28A8.6; CE18977.
DR InterPro; IPR006570; SPK.
DR Pfam; PF04435; SPK; 2.
DR SMART; SM00583; SPK; 2.
KW Hypothetical protein.
SQ SEQUENCE 627 AA; 71794 MW; 4A5654D3658EC1CC CRC64;

Query Match 73.5%; Score 36; DB 2; Length 627;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9
|:|:|:|
Db 62 KMNEYSIEL 70

RESULT 14
Q9LQW1 PRELIMINARY; PRT; 635 AA.
AC Q9LQW1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F10B6.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,
RA Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
RA Hansen N.F., Huizar L., Kremetskaia I., Lenz C., Li J., Liu S.,
RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,
RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006917; AAP79248.1; -.
DR InterPro; IPR007145; MAP65_ASE1.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 635 AA; 72664 MW; 2871C7DA1E4DA65E CRC64;

Query Match 73.5%; Score 36; DB 2; Length 635;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9
|:|:|:|
Db 169 KLINEYQTHL 177

RESULT 15
Q8DVD0 PRELIMINARY; PRT; 753 AA.
AC Q8DVD0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

DE ATP-dependent protease ClpE.  
GN Name=clpE; OrderedLocusNames=SMU.562;  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1309;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=UA159 / ATCC 700610 / Serotype c;  
RX MEDLINE=22395063; PubMed=12397186; DOI=10.1073/pnas.172501299;  
RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,  
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,  
RA Ferretti J.J.;  
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental  
pathogen.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
DR EMBL; AE014901; AAN58303.1; -.  
DR HSSP; P03815; 1JBK.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0004518; F:nuclease activity; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003959; AAA\_ATPase\_central.  
DR InterPro; IPR001270; Chaprinin\_clpA/B.  
DR InterPro; IPR001943; UvrB/C.  
DR Pfam; PF00004; AAA; 1.  
DR Pfam; PF02151; UVR; 1.  
DR PRINTS; PR00300; CLP/PROTEASEA.  
DR SMART; SM00382; AAA; 2.  
DR PROSITE; PS00870; CLPAB\_1; 1.  
DR PROSITE; PS00871; CLPAB\_2; 1.  
DR PROSITE; PS0151; UVR; 1.  
KW ATP-binding; Complete proteome; Protease.  
SQ SEQUENCE 753 AA; 83733 MW; EFA2A0D75FE71D5E CRC64;

Query Match 73.5%; Score 36; DB 2; Length 753;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9  
|:| |:  
Db 7 KLNEATHL 15

Search completed: July 13, 2005, 17:29:15  
Job time : 54.2222 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-31

Perfect score: 49

Sequence: 1 KMNEYTVHL 9

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCITUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	9	3	US-09-502-600-31
2	49	100.0	9	3	US-09-502-600-89
3	49	100.0	9	3	US-09-502-600-109
4	49	100.0	9	4	US-09-918-243-31
5	49	100.0	9	4	US-09-918-243-89
6	49	100.0	9	4	US-09-918-243-109
7	49	100.0	144	4	US-09-618-259-4
8	49	100.0	154	3	US-09-261-416-7
9	49	100.0	224	3	US-08-944-483-33
10	49	100.0	225	2	US-08-557-146-12
11	49	100.0	225	2	US-09-027-337-4
12	49	100.0	225	2	US-09-154-344-12
13	49	100.0	225	4	US-09-644-600-4
14	49	100.0	225	4	US-09-654-600A-4
15	49	100.0	253	2	US-08-557-146-2
16	49	100.0	253	2	US-08-824-874-3
17	49	100.0	253	2	US-09-154-344-2
18	49	100.0	253	3	US-08-930-188-2
19	49	100.0	253	3	US-09-210-084-3
20	49	100.0	253	4	US-09-764-762-3
21	49	100.0	253	5	PCR-US96-04294-2
22	49	100.0	265	4	US-09-949-016-7716
23	39	79.6	9	3	US-09-502-600-122
24	39	79.6	9	4	US-09-918-243-122
25	37	75.5	9	3	US-09-502-600-93
26	37	75.5	9	4	US-09-918-243-93
27	37	75.5	463	4	US-09-540-236-2942

28 35 71.4 415 4 US-09-949-016-11065 Sequence 11065, A  
29 33 67.3 204 3 US-09-323-872A-13 Sequence 13, Appl  
30 33 67.3 204 4 US-09-072-433-22 Sequence 22, Appl  
31 33 67.3 358 4 US-09-710-279-2136 Sequence 2136, Ap  
32 33 67.3 365 4 US-09-710-279-470 Sequence 470, App  
33 33 67.3 366 3 US-09-134-001C-5502 Sequence 5502, Ap  
34 33 67.3 799 3 US-08-909-954-4 Sequence 4, Appl  
35 33 67.3 804 3 US-08-909-954-2 Sequence 2, Appl  
36 32 65.3 60 3 US-09-134-001C-5184 Sequence 5184, Ap  
37 32 65.3 116 4 US-09-732-210-882 Sequence 882, App  
38 32 65.3 121 3 US-09-091-725-49 Sequence 49, Appl  
39 32 65.3 144 4 US-09-270-767-47539 Sequence 47539, A  
40 32 65.3 255 4 US-09-902-540-12375 Sequence 12375, A  
41 32 65.3 393 4 US-09-248-796A-26699 Sequence 26699, A  
42 32 65.3 698 4 US-09-943-016-10644 Sequence 10644, A  
43 32 65.3 752 4 US-09-583-110-2714 Sequence 2714, Ap  
44 32 65.3 755 4 US-09-107-433-4628 Sequence 4628, Ap  
45 32 65.3 792 4 US-09-134-000C-5895 Sequence 5895, Ap

#### ALIGNMENTS

RESULT 1  
US-09-502-600-31  
; Sequence 31, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 31  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-502-600-31

Query Match 100.0%; Score 49; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
| | | | | | | |  
Db 1 KMNEYTVHL 9

RESULT 2  
US-09-502-600-89  
; Sequence 89, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 89  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-502-600-89

Query Match 100.0%; Score 49; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
|||||

Db 1 KMNEYTVHL 9

## RESULT 3

US-09-502-600-109  
; Sequence 109, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE OF INVENTION: Ovarian Cancer  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR FILING DATE: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 109  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-502-600-109

Query Match 100.0%; Score 49; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
|||||

Db 1 KMNEYTVHL 9

## RESULT 4

US-09-918-243-31  
; Sequence 31, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 31  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-918-243-31

Query Match 100.0%; Score 49; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
|||||

Db 1 KMNEYTVHL 9

## RESULT 5

US-09-918-243-89  
; Sequence 89, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 89  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-918-243-89

Query Match 100.0%; Score 49; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
|||||

Db 1 KMNEYTVHL 9

## RESULT 6

US-09-918-243-109  
; Sequence 109, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 109  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-918-243-109

Query Match 100.0%; Score 49; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9  
|||||

Db 1 KMNEYTVHL 9

## RESULT 7

US-09-618-259-4  
; Sequence 4, Application US/09618259  
; Patent No. 6642013



```
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: NO. 6642013el Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/09/618,259
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURES:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-09-618-259-4

Query Match          100.0%; Score 49; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMEYTVHL 9
Db 9 KMEYTVHL 17

RESULT 8
US-09-261-416-7
; Sequence 7, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 7
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Unknown
; FEATURES:
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
; OTHER INFORMATION: chymotryptic enzyme (Scce) homologous to similar
; OTHER INFORMATION: domain in TAGD-12
US-09-261-416-7

Query Match          100.0%; Score 49; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMEYTVHL 9
Db 18 KMEYTVHL 26

RESULT 9
US-08-944-483-33
; Sequence 33, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.

; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-33

Query Match          100.0%; Score 49; DB 3; Length 224;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMEYTVHL 9
Db 43 KMEYTVHL 51

RESULT 10
US-08-557-146-12
; Sequence 12, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-557-146-12

Query Match 100.0%; Score 49; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKNQYTVHL 9
Db 44 KKNQYTVHL 52

RESULT 11
US-09-027-337-4
; Sequence 4, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotsoshi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in
; FILE OF INVENTION: Breast and Ovarian Carcinomas
; FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027,337B
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to
; OTHER INFORMATION: similar domain in TAGD-15
US-09-027-337-4

Query Match 100.0%; Score 49; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKNQYTVHL 9
Db 44 KKNQYTVHL 52

RESULT 12
US-09-154-344-12
; Sequence 12, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; FILE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
```

```
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-154-344-12

Query Match 100.0%; Score 49; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKNQYTVHL 9
Db 44 KKNQYTVHL 52

RESULT 13
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotsoshi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
; FILE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-644-600-4

Query Match 100.0%; Score 49; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKNQYTVHL 9
Db 44 KKNQYTVHL 52
```

RESULT 14  
US-09-654-600A-4  
; Sequence 4, Application US/09654600A  
; Patent No. 6649741  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotochi  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
; FILE REFERENCE: D6064CIP/D  
; CURRENT APPLICATION NUMBER: US/09/654,600A  
; CURRENT FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 09/421,213  
; 09/027,337  
; PRIOR FILING DATE: 1999-10-20  
; 1998-02-20  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: SCCE  
US-09-654-600A-4

Query Match 100.0%; Score 49; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNEYTVHL 9  
Db 44 KNEYTVHL 52

RESULT 15  
US-08-557-146-2  
; Sequence 2, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sterner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-557-146-2  
Query Match 100.0%; Score 49; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KNEYTVHL 9  
Db 72 KNEYTVHL 80  
Search completed: July 13, 2005, 17:34:24  
Job time : 20.4444 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.889 Seconds  
(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-31

Perfect score: 49

Sequence: 1 KMNEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pap.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pap.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pap.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pap.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pap.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pap.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pap.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pap.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pap.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	49	100.0	9	9 US-09-918-243-31	Sequence 31, Appl
2	49	100.0	9	9 US-09-918-243-89	Sequence 89, Appl
3	49	100.0	9	9 US-09-918-243-109	Sequence 109, Appl
4	49	100.0	9	9 US-09-905-083-31	Sequence 31, Appl
5	49	100.0	9	9 US-09-905-083-89	Sequence 89, Appl
6	49	100.0	9	9 US-09-905-083-109	Sequence 109, Appl
7	49	100.0	9	15 US-10-372-521-31	Sequence 31, Appl
8	49	100.0	9	15 US-10-372-521-89	Sequence 89, Appl
9	49	100.0	9	15 US-10-372-521-109	Sequence 109, Appl
10	49	100.0	9	15 US-10-831-075-31	Sequence 31, Appl
11	49	100.0	9	16 US-10-831-075-89	Sequence 89, Appl

12	49	100.0	9	15 US-10-831-075-109	Sequence 109, Appl
13	49	100.0	97	15 US-10-262-511-100	Sequence 100, Appl
14	49	100.0	144	9 US-09-796-294-4	Sequence 4, Appl
15	49	100.0	144	14 US-10-461-787-4	Sequence 4, Appl
16	49	100.0	181	15 US-10-262-511-98	Sequence 98, Appl
17	49	100.0	198	15 US-10-262-511-96	Sequence 96, Appl
18	49	100.0	224	11 US-09-789-210-33	Sequence 33, Appl
19	49	100.0	224	15 US-10-262-511-104	Sequence 104, Appl
20	49	100.0	225	15 US-10-600-187-4	Sequence 4, Appl
21	49	100.0	247	15 US-10-262-511-102	Sequence 102, Appl
22	49	100.0	250	15 US-10-262-511-92	Sequence 92, Appl
23	49	100.0	252	15 US-10-262-511-94	Sequence 94, Appl
24	49	100.0	253	9 US-09-888-615-98	Sequence 98, Appl
25	49	100.0	253	9 US-09-764-762-3	Sequence 3, Appl
26	49	100.0	253	14 US-10-071-214-2	Sequence 2, Appl
27	49	100.0	253	14 US-10-071-214-48	Sequence 48, Appl
28	49	100.0	253	14 US-10-264-283-90	Sequence 90, Appl
29	49	100.0	253	15 US-10-295-027-498	Sequence 498, Appl
30	49	100.0	253	15 US-10-173-999-48	Sequence 48, Appl
31	49	100.0	253	16 US-10-408-765A-639	Sequence 639, Appl
32	49	100.0	253	16 US-10-643-795A-95	Sequence 95, Appl
33	49	100.0	253	17 US-10-948-518-95	Sequence 95, Appl
34	49	100.0	253	17 US-10-868-490A-1	Sequence 1, Appl
35	49	100.0	257	15 US-10-344-394-38	Sequence 38, Appl
36	40	81.6	141	16 US-10-437-963-194726	Sequence 194726
37	40	81.6	226	14 US-10-071-214-49	Sequence 49, Appl
38	39	79.6	9	9 US-09-918-243-122	Sequence 122, Appl
39	39	79.6	9	9 US-09-905-083-122	Sequence 122, Appl
40	39	79.6	9	15 US-10-372-521-122	Sequence 122, Appl
41	39	79.6	9	16 US-10-831-075-122	Sequence 122, Appl
42	39	79.6	81	15 US-10-424-599-175183	Sequence 175183
43	37	75.5	9	9 US-09-918-243-93	Sequence 93, Appl
44	37	75.5	9	9 US-09-905-083-93	Sequence 93, Appl
45	37	75.5	9	15 US-10-372-521-93	Sequence 93, Appl

## ALIGNMENTS

### RESULT 1

US-09-918-243-31  
; Sequence 31, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 31  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-09-918-243-31

Query Match 100.0%; Score 49; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

Db 1 KMNEYTVHL 9

### RESULT 2

```
US-09-918-243-89
; Sequence 89, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-89

Query Match      100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KMNQYTVHL 9
Db      1 KMNQYTVHL 9

RESULT 3
US-09-918-243-109
; Sequence 109, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-109

Query Match      100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KMNQYTVHL 9
Db      1 KMNQYTVHL 9

RESULT 4
US-09-905-083-31
; Sequence 31, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/D/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-31

Query Match      100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KMNQYTVHL 9
Db      1 KMNQYTVHL 9

RESULT 5
US-09-905-083-89
; Sequence 89, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/D/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-89

Query Match      100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KMNQYTVHL 9
Db      1 KMNQYTVHL 9

RESULT 6
US-09-905-083-109
; Sequence 109, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/D/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-109

Query Match      100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KMNQYTVHL 9
Db      1 KMNQYTVHL 9
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-109

Query Match      100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 K0NEYTVHL 9
Db 1 K0NEYTVHL 9

RESULT 7
US-10-372-521-31
; Sequence 31, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-10-372-521-31

Query Match      100.0%; Score 49; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 K0NEYTVHL 9
Db 1 K0NEYTVHL 9

RESULT 8
US-10-372-521-89
; Sequence 89, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-10-372-521-89

Query Match      100.0%; Score 49; DB 15; Length 9;
```

```
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 K0NEYTVHL 9
Db 1 K0NEYTVHL 9

RESULT 9
US-10-372-521-109
; Sequence 109, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-10-372-521-109

Query Match      100.0%; Score 49; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 K0NEYTVHL 9
Db 1 K0NEYTVHL 9

RESULT 10
US-10-831-075-31
; Sequence 31, Application US/10831075
; Publication No. US20040224891A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP3
; CURRENT APPLICATION NUMBER: US/10/831,075
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 10/372,521
; PRIOR FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-10-831-075-31

Query Match      100.0%; Score 49; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 K0NEYTVHL 9
Db 1 K0NEYTVHL 9
```

## RESULT 11

US-10-831-075-89  
; Sequence 89, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 89  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-10-831-075-89

Query Match 100.0%; Score 49; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9  
| | | | | | | | | |  
DB 1 KNEYTVHL 9

## RESULT 12

US-10-831-075-109  
; Sequence 109, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 109  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 72-80 of the SCCE protein  
US-10-831-075-109

Query Match 100.0%; Score 49; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9  
| | | | | | | | | |  
DB 1 KNEYTVHL 9

## RESULT 13

US-10-262-511-100  
; Sequence 100, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda

; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 100  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-100

Query Match 100.0%; Score 49; DB 15; Length 97;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9  
| | | | | | | | | |  
DB 53 KNEYTVHL 61

## RESULT 14

US-09-796-294-4  
; Sequence 4, Application US/09796294



```

; Patent No. US20020037581A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Extracellular Serine Protease
; FILE REFERENCE: D6020CIP3
; CURRENT APPLICATION NUMBER: US/09/796,294
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-09-796-294-4

```

```

Query Match      100.0%; Score 49; DB 9; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KNEYTVHL 9
Db 9 KNEYTVHL 17

```

```

RESULT 15
US-10-461-787-4
; Sequence 4, Application US/10461787
; Publication No. US20030199010A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: No. US20030199010A1el Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/10/461,787
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US/09/618,259
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-10-461-787-4

```

```

Query Match      100.0%; Score 49; DB 14; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KNEYTVHL 9
Db 9 KNEYTVHL 17

```

Search completed: July 13, 2005, 18:23:28  
Job time : 67.8889 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083A-32

Perfect score: 40

Sequence: 1 RLSSWVKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A: Genesep16Dec04: \*  
1: genesep1980s: \*  
2: genesep1990s: \*  
3: genesep2000s: \*  
4: genesep2001s: \*  
5: genesep2002s: \*  
6: genesep2003as: \*  
7: genesep2003bs: \*  
8: genesep2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40	100.0	9	4	AAE08237	Human str
2	40	100.0	9	8	ADR68793	Human str
3	40	100.0	144	8	ADI39727	Stratum c
4	40	100.0	144	8	ADI37151	Stratum c
5	40	100.0	224	6	ADA05744	Human NOV
6	40	100.0	224	8	ADN62908	Human NOV
7	40	100.0	225	4	AB988502	Human Str
8	40	100.0	247	6	ADA05742	Human NOV
9	40	100.0	247	8	ADN62906	Human NOV
10	40	100.0	250	6	ADA05732	Human NOV
11	40	100.0	250	8	ADN62896	Human NOV
12	40	100.0	252	6	ADA05734	Human NOV
13	40	100.0	252	8	ADN62898	Human NOV
14	40	100.0	253	2	AAR67888	Human str
15	40	100.0	253	2	AAW05383	Human amy
16	40	100.0	253	5	ABB84421	Human SCC
17	40	100.0	253	5	ABB84406	Human SCC
18	40	100.0	253	5	AAU82740	Amino aci
19	40	100.0	253	6	ABU07440	Protein d
20	40	100.0	253	6	ABU07471	Protein d
21	40	100.0	253	6	ABR58471	Human str
22	40	100.0	253	7	ADB80484	Ovarian c
23	40	100.0	253	7	ADJ68833	Human hea
24	40	100.0	253	7	ADN39180	Cancer/an
25	40	100.0	253	8	ADL06515	Human tum

26	40	100.0	253	8	ADN04182	Antipsori
27	40	100.0	253	8	ADR72880	Human ova
28	40	100.0	257	3	AB21326	Human HSC
29	36	90.0	9	4	AAE08314	Human str
30	36	90.0	9	8	ADR68871	Human str
31	33	82.5	394	5	ABP52834	Anopheles
32	33	82.5	394	7	ABR84477	Mosquito
33	33	82.5	417	7	ADD15250	African m
34	32	80.0	589	6	ABU48685	Protein e
35	31	77.5	9	4	AAE08292	Human str
36	31	77.5	9	4	AAE08284	Human str
37	31	77.5	9	8	ADR68840	Human str
38	31	77.5	9	8	ADR68848	Human str
39	31	77.5	136	4	ASG23378	Novel hum
40	31	77.5	229	3	AAG19496	Arabidops
41	31	77.5	229	3	AAG52872	Arabidops
42	31	77.5	243	5	ABB84419	Bovine SC
43	31	77.5	249	5	ABB84420	Porcine S
44	31	77.5	390	3	AAG19495	Arabidops
45	31	77.5	390	3	AAG52871	Arabidops

## ALIGNMENTS

## RESULT 1

AAE08237  
ID AAE08237 standard; peptide; 9 AA.

XX AAE08237;

XX DT 01-NOV-2001 (first entry)

XX DE Human stratum corneum chymotrypsin enzyme peptide #2 (residues 123-131).

XX KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;  
XX KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;  
XX KW antisense therapy; malignant hyperplasia.

XX OS Homo sapiens.

XX PN WO200159158-A1.

XX PD 16-AUG-2001.

XX PF 07-FEB-2001; 2001WO-US003977.

XX PR 11-FEB-2000; 2000US-00502600.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'brien TU;

XX DR WPI; 2001-514676/56.

XX PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX PS Claim 25; Page 102; 127pp; English.

XX CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

SQ Sequence 9 AA;

Query Match 100.0%; Score 40; DB 4; Length 9;

```
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
Db 1 RLSSMWKVKV 9
|||||

RESULT 2
ID ADR68793 standard; peptide; 9 AA.
XX AC ADR68793;
XX DT 02-DEC-2004 (first entry)
XX DE Human stratum corneum chymotryptic enzyme peptide fragment SEQ ID NO:32.
XX KW serine protease; stratum corneum chymotryptic enzyme; SCCE;
XX KW immune response; ovarian cancer; lung cancer; prostate cancer;
XX KW pancreatic cancer; colon cancer.
XX OS Homo sapiens.
XX PN WO2004075723-A2.
XX PD 10-SEP-2004.
XX PF 20-FEB-2004; 2004WO-US005134.
XX PR 21-FEB-2003; 2003US-00372521.
XX PA (UYAR-) UNIV ARKANSAS.
XX PI O'brien TJ, Cannon MJ, Santin A;
XX DR WPI; 2004-653294/63.
XX PT Use of stratum corneum chymotryptic enzyme (SCCE) peptides, for
PT vaccinating an individual against SCCE, and in monitoring and developing
PT immunotherapies for ovarian and other malignancies.
XX PS Claim 5; SEQ ID NO 32; 117pp; English.
XX CC The invention relates to a novel method for vaccinating an individual
CC against stratum corneum chymotryptic enzyme (SCCE), comprising inoculating
CC an individual with a SCCE peptide, which elicits an immune response in
CC the individual. A peptide of the invention acts as a stratum corneum
CC chymotryptic enzyme inhibitor. The SCCE peptide is useful for vaccinating
CC an individual against SCCE, particularly an individual having, suspected
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon
CC cancer. The oligonucleotide is useful for treating a neoplastic state in
CC an individual, such as ovarian, breast, lung, colon, prostate, or
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The
CC peptides are also useful in the monitoring and development of
CC immunotherapies for ovarian and other malignancies. The present sequence
CC represents a peptide fragment of serine protease SCCE (stratum corneum
CC chymotryptic enzyme).
XX SQ Sequence 9 AA;
Query Match 100.0%; Score 40; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
Db 1 RLSSMWKVKV 9
|||||

RESULT 3
AD139727
ID AD139727 standard; protein; 144 AA.
XX AC AD139727;
XX DT 15-APR-2004 (first entry)
XX DE Stratum corneum chymotryptic enzyme (scce) catalytic domain.
XX KW Immune T cell; dendritic cell; extracellular serine protease;
XX KW tumour antigen derived gene-14; TADG-14; carcinoma;
XX KW stratum corneum chymotryptic enzyme; scce.
XX OS Unidentified.
XX PN US6642013-B1.
XX PD 04-NOV-2003.
XX PF 18-JUL-2000; 2000US-00618259.
XX PR 21-AUG-1997; 97US-00915659.
XX PR 21-AUG-1998; 98US-00137944.
XX PA (UYAR-) UNIV ARKANSAS MEDICAL SCI.
XX PI O'brien TJ, Underwood LJ;
XX DR WPI; 2004-118109/12.
XX PF Production of activated immune cells or dendritic cells by exposing
PT immune cells to tumor antigen derived gene protein fragment consisting of
PT amino acid sequences.
XX PS Example 1; SEQ ID NO 4; 44pp; English.
XX CC The present invention relates to novel activated immune T cells or
CC dendritic cells directed toward extracellular serine protease termed
CC tumour antigen derived gene-14 (TADG-14). The method of the invention
CC involves exposing the immune cells to a TADG-14 protein fragment, where
CC exposure to the TADG-14 protein fragment activates the immune cells. The
CC invention is used for the production of activated immune T cells or
CC dendritic cells. The invention allows screening to identify proteases
CC overexpressed in carcinoma. The present sequence is stratum corneum
CC chymotryptic enzyme (scce) catalytic domain. This sequence is used in the
CC invention.
XX SQ Sequence 144 AA;
Query Match 100.0%; Score 40; DB 8; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
Db 60 RLSSMWKVKV 68
|||||

RESULT 4
AD137151
ID AD137151 standard; protein; 144 AA.
XX AC AD137151;
XX DT 06-MAY-2004 (first entry)
XX DE Stratum corneum chymotryptic enzyme (scce) catalytic domain.
XX KW Serine protease; tumour antigen derived gene-14; TADG-14;
XX KW neoplastic state; cancer; ovary; breast; lung; colon; prostate;
XX KW chymotryptic enzyme; scce; enzyme.
XX OS Homo sapiens.
XX PN US2003199010-A1.
```

XX 23-OCT-2003.  
 XX 13-JUN-2003; 2003US-00461787.  
 PF 21-AUG-1997; 97US-00915659.  
 PR 21-AUG-1998; 98US-00137944.  
 PR 18-JUL-2000; 2000US-00618259.  
 XX (UYAR-) UNIV ARKANSAS.  
 PA O'brien TJ, Underwood LJ;  
 PI WPI; 2004-141550/14.  
 DR Novel isolated DNA encoding tumor antigen derived gene-14 protein, useful  
 XX for treating neoplastic state (such as ovarian cancer, breast cancer,  
 PT lung cancer, colon cancer, prostate cancer) in an individual.  
 PT Example 1; SEQ ID NO 4; 46pp; English.  
 PS The invention relates to extracellular serine protease termed tumour  
 XX antigen derived gene-14 (TAGD-14) and its nucleic acid. Composition  
 CC comprising TAGD-14 peptide is useful for treating a neoplastic state in  
 CC an individual. The neoplastic state is chosen from ovarian cancer, breast  
 CC cancer, lung cancer, colon cancer, prostate cancer in which TAGD-14 is  
 CC overexpressed. The present sequence is Stratum corneum chymotryptic  
 CC enzyme (sccc) catalytic domain. This sequence is used in the  
 CC exemplification of the invention.  
 XX Sequence 144 AA;  
 SQ

Query Match 100.0%; Score 40; DB 8; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RLSSMWKVV 9  
 Db 60 RLSSMWKVV 69  
 |||||

RESULT 5  
 ADA05744  
 ID ADA05744 standard; protein; 224 AA.  
 AC ADA05744;  
 AC  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV18g protein SEQ ID NO:104.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003029424-A2.  
 PN  
 XX 10-APR-2003.  
 PD  
 XX 02-OCT-2002; 2002WO-US031373.  
 PF  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 09-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 01-OCT-2002; 2002US-00262511.  
 XX (CURA-) CURAGEN CORP.  
 PA Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX WPI; 2003-381626/36.  
 DR N-PSDB; ADA05743.  
 XX  
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 172; 586pp; English.  
 XX  
 CC The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.

XX  
SQ Sequence 224 AA;  
Query Match 100.0%; Score 40; DB 6; Length 224;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
|||||  
Db 109 RLSSMVKKV 117

## RESULT 6

ADN62908  
ID ADN62908 standard; protein; 224 AA.

XX

AC ADN62908;

DT 01-JUL-2004 (first entry)

DE Human NOV18g.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.

XX Homo sapiens.

OS  
XX US2004038223-A1.

PN  
XX 26-FEB-2004.

PD  
XX 01-OCT-2002; 2002US-00262511.

PF  
XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.  
PA (JUJ/) JU J.  
PA (LIL/) LI L.  
PA (GUOX/) GUO X.  
PA (PAT/) PATTURAJAN M.  
PA (SPYT/) SPYTEK K A.  
PA (EDIN/) EDINGER S R.  
PA (ELLE/) ELLERMAN K.  
PA (MALY/) MALYANKAR U M.  
PA (ORTT/) ORT T.  
PA (GORM/) GORMAN L.  
PA (ZERH/) ZERHUSEN B D.  
PA (ANDE/) ANDERSON D W.  
PA (ZHON/) ZHONG M.  
PA (CATT/) CATTERTON E.  
PA (JIMW/) JI W.  
PA (MILL/) MILLER C E.  
PA (RAST/) RASTELLI L.  
PA (STON/) STONE D J.  
PA (PENA/) PENNA C E A.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.  
PA (ROTH/) ROTHENBERG M E.  
PA (LEAC/) LEACH M D.  
PA (AGEE/) AGE E M L.  
PA (BERG/) BERGHS C.  
PA (DIP/) DIPIPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CEA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX MPI; 2004-213931/20.  
DR N-FSDB; ADN62907.

Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

Claim 1; SEQ ID NO 104; 395pp; English.

The invention relates to isolated NOVX polypeptides and polynucleotides.  
NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
treat a medical condition in human related to the aberrant expression and  
activity of NOVX polypeptides. For example, NOVX polypeptides and  
polynucleotides may be used to treat disorders associated with decreased  
expression or activity of NOVX by supplementing the patient our  
production or to rectify mutations. Conversely, antisense NA molecules  
may be administered to down regulate expression of NOVX polypeptides by  
binding with the cells own genes and preventing their expression. NOVX  
polynucleotides and complementary sequences may also be used as DNA  
probes in diagnostic assays to detect and quantitate the presence of  
similar sequences in samples, and so which patients may be in need of  
restorative therapy. NOVX polypeptides may also be used as antigens in  
the production of antibodies and in assays to identify modulators  
(agonists and antagonists) of the expression and activity of NOVX. The  
anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
used to modulate NOVX polynucleotide expression and activity of NOVX  
polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
diagnostic agents for detecting the presence of NOVX in samples. NOVX  
polypeptides and polynucleotides may be used in this way to prevent,  
diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
haematopoietic disorders, and the various dyslipidaemias, metabolic  
disturbances associated with obesity, the metabolic syndrome X and  
wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 224 AA;

Query Match 100.0%; Score 40; DB 8; Length 224;  
Best Local Similarity 100.0%; Pred. No. 2.7; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 RLSSMVKKV 9

|||||

Db 109 RLSSMVKKV 117

#### RESULT 7

AA098502  
ID AAB98502 standard; protein; 225 AA.

XX AAB98502;

DT 03-AUG-2001 (first entry)

XX Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.

XX Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;

KW tumour antigen-derived gene 15; serine protease;

KW Stratum Corneum Chymotryptic Enzyme; SCCE.

XX Homo sapiens.

OS WO200129056-A1.

PN 26-APR-2001.

XX 20-OCT-2000; 2000WO-US029095.

PF 20-OCT-1999; 99US-00421213.

PR (UYAR-) UNIV ARKANSAS.

PA O'brien TJ, Tanimoto H;

PI WPI; 2001-381031/40.

DR Novel extracellular serine protease, termed tumor antigen-derived gene 15  
PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,  
PT treatment, prevention of cancer, particularly breast, ovarian cancer.

PS Example 10; Fig 1; 130pp; English.

XX The present invention relates to human tumour antigen-derived gene 15  
CC (TADG-15) protein and coding sequence (see AAB98501 and AAB98500). TADG-  
CC 15 is an extracellular serine protease. It was found that TADG-15 is over  
CC-expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20  
CC residues that lack TADG-15 protease activity are useful for vaccinating  
CC an individual against TADG-15, having, suspected of having or at risk of  
CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic  
CC or therapeutic target in cancer. The present sequence was used in a  
CC sequence homology alignment with the catalytic domain of TADG-15

XX Sequence 225 AA;

Query Match 100.0%; Score 40; DB 4; Length 225;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

|||||

Db 95 RLSSMVKKV 103

#### RESULT 8

ADA05742

ID ADA05742 standard; protein; 247 AA.

XX ADA05742;

DT 06-NOV-2003 (first entry)

XX Human NOV18f protein SEQ ID NO:102.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytostatic; nootropic; neuroprotective;

KW antiparkinsonian; antilipemic; gene therapy; human disease;

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

OS WO2003029424-A2.

PN 10-APR-2003.

PF 02-OCT-2002; 2002WO-US031373.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 12-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CBA, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eiben AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

DR N-PSDB; ADA05741.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.

PS Claim 1; Page 172; 586pp; English.

CC The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipaeamic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 247 AA;

Query Match 100.0%; Score 40; DB 6; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 109 RLSSMVKKV 117

RESULT 9

ADN62906

ID ADN62906 standard; protein; 247 AA.

AC ADN62906;

DT 01-JUL-2004 (first entry)

DE Human NOV18f.

KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

OS Homo sapiens.

XX US2004038223-A1.

PN 26-FEB-2004.

PD 01-OCT-2002; 2002US-00262511.

PF 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.

PA (JUJJ/) JU J.

PA (LILL/) LI L.

PA (GUOX/) GUO X.

PA (PATT/) PATTURAJAN M.

PA (SPYT/) SPYTEK K A.

PA (EDIN/) EDINGER S R.

PA (ELLE/) ELLERMAN K.

PA (MALY/) MALYANKAR U M.

PA (ORTT/) ORT T.

PA (GORM/) GORMAN L.

PA (ZERH/) ZERHUSEN B D.

PA (ANDE/) ANDERSON D W.

PA (ZHON/) ZHONG M.

PA (CATT/) CATTERTON E.

PA (JIWW/) JI W.

PA (MILL/) MILLER C E.

PA (RST/) RASTELLI L.

PA (STON/) STONE D J.

PA (PENA/) PENNA C E A.

PA (SHEN/) SHENOY S G.

PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.

PA (LEAC/) LEACH M D.

PA (AGEE/) AGEE M L.

PA (BERG/) BERGHS C.

PA (DIPI/) DIPIPO V A.

PA (EISE/) EISEN A.

PA (GANG/) GANGOLLI E A.

PA (RIEG/) RIEGER D K.

PA (SPAD/) SPADERNA S K.

XX

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Bergths C, Dipippo VA;

PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI, 2004-213931/20.

DR N-FSDB; ADN62905.

XX

PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,





CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX  
SQ Sequence 250 AA;

Query Match 100.0%; Score 40; DB 6; Length 250;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RLSSMVKKV 9  
|||||||  
Db 120 RLSSMVKKV 128

RESULT 11  
ADN62896  
ID ADN62896 standard; protein; 250 AA.  
XX  
AC ADN62896;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human NOV18a.  
XX  
KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2004038223-A1.  
XX  
PD 26-FEB-2004.  
XX  
PF 01-OCT-2002; 2002US-00262511.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 15-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 28-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
XX

PA (SMIT//) SMITHSON G.  
PA (MILL//) MILLET I.  
PA (PEYM//) PEYMAN J A.  
PA (KEKU//) KEKUDA R.  
PA (JUJJ//) JU J.  
PA (LILL//) LI L.  
PA (GUOX//) GUO X.  
PA (PATT//) PATTURAJAN M.  
PA (SPYT//) SPYTEK K A.  
PA (EDIN//) EDINGER S R.  
PA (ELLE//) ELLERMAN K.  
PA (MALY//) MALYANKAR U M.  
PA (ORTT//) ORT T.  
PA (GORM//) GORMAN L.  
PA (ZERH//) ZERHUSEN B D.  
PA (ANDE//) ANDERSON D W.  
PA (ZHON//) ZHONG M.  
PA (CATT//) CATTERTON E.  
PA (JIWW//) JI W.  
PA (MILL//) MILLER C E.  
PA (RAST//) RASTELLI L.  
PA (STON//) STONE D J.  
PA (PENA//) PENNA C E A.  
PA (SHEN//) SHENOY S G.  
PA (SHIM//) SHIMKETS R A.  
PA (ROTH//) ROTHENBERG M E.  
PA (LEAC//) LEACH M D.  
PA (AGEE//) AGEE M L.  
PA (BERG//) BERGHS C.  
PA (DIPJ//) DIPPIO V A.  
PA (EISE//) EISEN A.  
PA (GANG//) GANGOLLI E A.  
PA (RIEG//) RIEGER D K.  
PA (SPAD//) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-213931/20.  
DR N-PSDB; ADN62895.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
PS Claim 1; SEQ ID NO 92; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent.  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,

CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX  
 SQ Sequence 250 AA;

Query Match 100.0%; Score 40; DB 8; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9  
 |||||  
 Db 120 RLSSMWKKV 128

RESULT 12  
 ADA05734  
 ID ADA05734 standard; protein; 252 AA.

AC ADA05734;

DT 06-NOV-2003 (first entry)

DE Human NOV18b protein SEQ ID NO:94.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

PN W02003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327433P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328058P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339268P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373815P.

(CURA-) CURAGEN CORP.

PA Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

DR N-PSDB; ADA05733.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.

PS Claim 1; Page 170; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX Sequence 252 AA;

Query Match 100.0%; Score 40; DB 6; Length 252;

Best Local Similarity 100.0%; Pred. No. 3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9

Db 122 RLSSMWKKV 130

RESULT 13

ADN62898

ID ADN62898 standard; protein; 252 AA.

XX ADN62898;

XX 01-JUL-2004 (first entry)

XX

DE Human NOV18b.  
XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2004038223-A1.  
XX  
PD 26-FEB-2004.  
XX  
PF 01-OCT-2002; 2002US-00262511.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0332668P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381644P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
XX (SMIT/) SMITHSON G.  
PA (MILL/) MILLET I.  
PA (PEYM/) PEYMAN J A.  
PA (KEKU/) KEKUDA R.  
PA (JULJ/) JU J.  
PA (LILL/) LI L.  
PA (GUOX/) GUO X.  
PA (PATT/) PATTURAJAN M.  
PA (SPYT/) SPYTEK K A.  
PA (EDIN/) EDINGER S R.  
PA (ELLE/) ELLERMAN K.  
PA (MALY/) MALYANKAR U M.  
PA (ORTT/) ORT T.  
PA (GORM/) GORMAN L.  
PA (ZERH/) ZERHUSEN B D.  
PA (ANDE/) ANDERSON D W.  
PA (ZHON/) ZHONG M.  
PA (CATT/) CATTERTON E.  
PA (JIWW/) JI W.  
PA (MILL/) MILLER C E.  
PA (RAST/) RASTELLI L.  
PA (STON/) STONE D J.  
PA (PENA/) PENNA C E A.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.  
PA (ROTH/) ROTHENBERG M E.  
PA (LEAC/) LEACH M D.

PA (AGEE/) AGEE M L.  
PA (BERG/) BERGHS C.  
PA (DIPI/) DIPIPPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-213931/20.  
DR N-FSDB; ADN62897.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
PS Claim 1; SEQ ID NO 94; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
XX represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 252 AA;  
Query Match 100.0%; Score 40; DB 8; Length 252;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RLSSMVKKV 9  
DB 122 RLSSMVKKV 130  
RESULT 14  
AAR67888  
ID AAR67888 standard; protein; 253 AA.  
XX  
AC AAR67888;  
XX  
DT 25-MAR-2003 (revised)  
DT 09-AUG-1995 (first entry)  
XX  
DE Human stratum corneum chymotrophic recombinant enzyme (SCCE).

XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
 KW callosities; keratosis pilaris; ichthyoses; eczema.  
 XX Homo sapiens.  
 XX WO9500651-A1.  
 XX 05-JAN-1995.  
 XX 20-JUN-1994; 94WO-IB000166.  
 XX 18-JUN-1993; 93DK-00000725.  
 XX (SYMB-) SYMBICOM AB.  
 XX Egelrud T, Hansson L;  
 XX WPI: 1995-052088/07.  
 XX N-PSDB; AAQ81203.  
 XX Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
 PT related vectors, transformed cells and polypeptides, useful for treating  
 PT skin disorders, e.g. acne or psoriasis, and for identification of  
 PT specific inhibitors.  
 XX Disclosure; Page 97; 137pp; English.  
 XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
 CC and skin care products, especially to treat and prevent acne, xeroderma,  
 CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
 CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
 CC recombinantly following mammal, insect, plant, or microorganism  
 CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX Sequence 253 AA;  
 SQ  
 Query Match 100.0%; Score 40; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RLSSMVKKV 9  
 DB 123 RLSSMVKKV 131  
 Search completed: July 13, 2005, 17:19:50  
 Job time : 77.6667 secs

---

XX WPI: 1996-464694/46.  
 DR N-PSDB; AAT39783.  
 XX New isolated human amyloid precursor protein protease - used to develop  
 PT prods. for the treatment or diagnosis of associated conditions, esp.  
 PT Alzheimer's disease.  
 XX Claim 1; Page 44-45; 55pp; English.  
 XX Human amyloid precursor protein protease (AAW05383) is involved in the  
 CC processing or clearance of amyloid precursor protein to form beta-amyloid  
 CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)  
 CC obtd. from a human lung library. Recombinant protease can be produced in  
 CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic  
 CC (partic. AV-120 host cells. It is used to develop products for the design  
 CC and testing of cpds. useful for treating or preventing conditions  
 CC associated with beta-amyloid peptide, esp. Alzheimer's disease  
 XX Sequence 253 AA;  
 SQ  
 Query Match 100.0%; Score 40; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RLSSMVKKV 9  
 DB 123 RLSSMVKKV 131  
 Search completed: July 13, 2005, 17:19:50  
 Job time : 77.6667 secs

---

XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
 KW callosities; keratosis pilaris; ichthyoses; eczema.  
 XX Homo sapiens.  
 XX WO9500651-A1.  
 XX 05-JAN-1995.  
 XX 20-JUN-1994; 94WO-IB000166.  
 XX 18-JUN-1993; 93DK-00000725.  
 XX (SYMB-) SYMBICOM AB.  
 XX Egelrud T, Hansson L;  
 XX WPI: 1995-052088/07.  
 XX N-PSDB; AAQ81203.  
 XX Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
 PT related vectors, transformed cells and polypeptides, useful for treating  
 PT skin disorders, e.g. acne or psoriasis, and for identification of  
 PT specific inhibitors.  
 XX Disclosure; Page 97; 137pp; English.  
 XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
 CC and skin care products, especially to treat and prevent acne, xeroderma,  
 CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
 CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
 CC recombinantly following mammal, insect, plant, or microorganism  
 CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX Sequence 253 AA;  
 SQ  
 Query Match 100.0%; Score 40; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RLSSMVKKV 9  
 DB 123 RLSSMVKKV 131  
 Search completed: July 13, 2005, 17:19:50  
 Job time : 77.6667 secs

---

RESULT 15  
 AAW05383  
 ID AAW05383 standard; protein; 253 AA.  
 XX AAW05383;  
 AC  
 XX 31-DEC-1996 (first entry)  
 DT  
 XX Human amyloid precursor protein protease.  
 DE  
 XX Amyloid precursor protein protease; Alzheimer's disease; diagnosis;  
 KW therapy.  
 KW Homo sapiens.  
 XX WO9631122-A1.  
 PN 10-OCT-1996.  
 PD  
 XX 02-APR-1996; 96WO-US004294.  
 PF  
 XX 04-APR-1995; 95US-00416257.  
 PR  
 XX (ELIL ) LILLY & CO ELI.  
 PA  
 XX Dixon EP, Johnstone EM, Little SP;  
 PI

**This page blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083a-32

Perfect score: 40

Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	253	2 A53968	serine proteinase
2	33	82.5	282	2 T35294	probable endo alph
3	33	82.5	372	2 S20056	para-hydroxybenzoa
4	33	82.5	560	2 S50401	hypothetical prote
5	32	80.0	350	1 C70009	ABC transporter (l
6	32	80.0	351	2 D90264	biotin synthase (b
7	32	80.0	358	2 T38914	para-hydroxybenzoa
8	32	80.0	589	2 A71277	arginine-trna liga
9	32	80.0	664	2 T50316	hypothetical Armad
10	31	77.5	912	2 H85470	hypothetical prote
11	31	77.5	925	2 T05012	hypothetical prote
12	30	75.0	239	2 T02473	hypothetical prote
13	30	75.0	240	2 T32363	hypothetical prote
14	30	75.0	318	2 I56519	taipoxin-associate
15	30	75.0	319	2 JC5402	vitamin D receptor
16	30	75.0	373	2 G75073	hypothetical prote
17	30	75.0	457	2 D85429	cytochrome P450 11
18	30	75.0	532	2 T01759	glycine hydroxymet
19	30	75.0	647	2 A49218	hemagglutinin homo
20	30	75.0	833	2 A47528	transcription fact
21	30	75.0	986	2 E90220	isoleucine-tRNA sy
22	30	75.0	1167	2 A82543	chromosome segrega
23	30	75.0	1846	2 T33079	hypothetical prote
24	29	72.5	253	2 JC5511	TATA-binding prote
25	29	72.5	285	2 A97255	S-adenosylmethioni
26	29	72.5	336	2 E69214	GTP-binding protei
27	29	72.5	337	2 AH2591	membrane lipoprote
28	29	72.5	337	2 B97374	Deinococcus radiol
29	29	72.5	345	2 F90194	threonine synthase

30	29	72.5	349	2 P84246	hypothetical prote
31	29	72.5	357	2 D96986	probable lipoprote
32	29	72.5	379	2 H75318	membrane lipoprote
33	29	72.5	553	2 A72228	hypothetical prote
34	29	72.5	571	2 S65060	phytoene desaturas
35	29	72.5	578	2 H82872	hypothetical prote
36	29	72.5	582	2 S29314	phytoene dehydroge
37	29	72.5	583	2 A45381	phytoene dehydroge
38	29	72.5	656	2 T23338	hypothetical prote
39	29	72.5	691	2 E70906	probable beta-gluc
40	29	72.5	839	2 S62963	hypothetical prote
41	29	72.5	1360	2 F96596	hypothetical prote
42	29	72.5	4077	2 T17484	hypothetical prote
43	29	72.5	4589	2 T14914	dynein beta heavy
44	28	70.0	41	2 G71376	hypothetical prote
45	28	70.0	63	2 AH2565	hypothetical prote

ALIGNMENTS

RESULT 1

A53968

serine proteinase SCCE precursor - human

N:Alternate names: stratum corneum chymotryptic enzyme

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A53968

R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A:Reference number: A53968; MUID:94308225; PMID:8034709

A:Accession: A53968

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <HAN>

A:Cross-references: UNIPROT:P49862; GB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G532504

C:Genetics: A:Gene: GDB:PRSS6; SCCE

A:Cross-references: GDB:377730

A:Map position: 7q35-7q35

C:Superfamily: trypsin; trypsin homology

F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9

Db 123 RLSSMVKKV 131

RESULT 2

T35294

probable endo alpha-1,4 polygalactosaminidase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C:Accession: T35294

R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21574

A:Accession: T35294

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-282 <SEE>

A:Cross-references: UNIPROT:Q9S2Q8; EMBL:AL096872; PIDN:CAB51262.1; GSPDB:GN00070; SCOEID

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEIDB:SCSF7.23c

Query Match 82.5%; Score 33; DB 2; Length 282;  
Best Local Similarity 87.5%; Pred. No. 14;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKK 8  
|||:||||  
Db 265 RLSSMLKK 272

RESULT 3  
S20056  
para-hydroxybenzoate polyprenyltransferase (EC 2.5.1.-) precursor - yeast (Saccharomyces  
N:Alternate names: 4-hydroxybenzoate hexaprenyltransferase; protein N3419; protein YNR04  
C:Species: Saccharomyces cerevisiae  
C:Date: 23-Apr-1993 #sequence\_revision 23-Apr-1993 #text\_change 09-Jul-2004  
C:Accession: S20056; S63372; JC2317  
R:Ashby, M.N.; Kutsuna, S.Y.; Ackerman, S.; Tzagoloff, A.; Edwards, P.A.  
J. Biol. Chem. 267, 4128-4136, 1992  
A:Title: COQ2 is a candidate for the structural gene encoding para-hydroxybenzoate:polyp  
A:Reference number: S20056; MUID:92156158; PMID:1740455  
A:Accession: S20056  
A:Molecule type: DNA  
A:Residues: 1-372 <ASH>  
A:Cross-references: UNIPROT:P32378; EMBL:M81698; NID:gl71252; PIDN:AAA34507.1; PID:gl712  
R:Pohl, T.M.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S63346  
A:Accession: S63372  
A:Molecule type: DNA  
A:Residues: 1-372 <POH>  
A:Cross-references: EMBL:271656; NID:gl302546; PID:e239832; PID:gl302547; MIPS:YNR041C  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:COQ2  
A:Cross-references: SGD:S0005324; MIPS:YNR041C  
A:Map position: 14R  
C:Keywords: isoprenoid biosynthesis; mitochondrion; transferase; transmembrane protein  
F:1-22/Domain: transit peptide [mitochondrion] #status predicted <TR>  
F:117-133/Domain: transmembrane #status predicted <TM1>  
F:172-188/Domain: transmembrane #status predicted <TM2>  
F:203-225/Region: polyprenyl diphosphate binding #status predicted

Query Match 82.5%; Score 33; DB 2; Length 372;  
Best Local Similarity 77.8%; Pred. No. 18;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9  
|||:||||  
Db 326 RLFSMLKKV 334

RESULT 4  
S50401  
hypothetical protein YMR145c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YMR375.114c  
C:Species: Saccharomyces cerevisiae  
C:Date: 13-Jan-1995 #sequence\_revision 24-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S50401  
R:Badcock, K.; Church, C.  
submitted to the EMBL Data Library, December 1994  
A:Reference number: S50388  
A:Accession: S50401  
A:Molecule type: DNA  
A:Residues: 1-560 <BAD>  
A:Cross-references: UNIPROT:P40215; EMBL:247071; NID:g606429; PIDN:CAA87359.1; PID:g6064  
C:Genetics:  
A:Gene: MIPS:YMR145c  
A:Cross-references: SGD:S0004753  
A:Map position: 13R  
C:Superfamily: NADH dehydrogenase

Query Match 82.5%; Score 33; DB 2; Length 560;  
Best Local Similarity 77.8%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9  
|||:||||  
Db 347 RLKTMVKKV 355

RESULT 5  
C70009  
ABC transporter (lipoprotein) homolog yufN - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: C70009  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C:Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi  
A: Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc  
A:Ich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,  
Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: C70009  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-350 <KUN>  
A:Cross-references: UNIPROT:O05252; GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15143.1  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yufN  
C:Superfamily: ABC transporter yufN

Query Match 80.0%; Score 32; DB 1; Length 350;  
Best Local Similarity 87.5%; Pred. No. 29;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSMVKKV 9  
|||:||||  
Db 269 LTMVKKV 276

RESULT 6  
D90264  
biotin synthase (bioB) [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: D90264  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-y  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: D90264  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-351 <KUR>  
A:Cross-references: UNIPROT:Q97226; GB:AE006641; NID:gl3814305; PIDN:AAK41371.1; GSPDB:G  
C:Genetics:  
A:Gene: bioB

Query Match 80.0%; Score 32; DB 2; Length 351;  
Best Local Similarity 66.7%; Pred. No. 29;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9  
|||:||||  
Db 341 RLDSLKKV 349



## RESULT 7

T38914  
para-hydroxybenzoate-polyphenyltransferase (EC 2.5.1.-) precursor, mitochondrial - fission  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T38914  
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z21817  
A:Accession: T38914  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-358 <PE>  
A:Cross-references: UNIPROT:Q10252; EMBL:Z69728; PIDN:CAA93575.1; GSPDB:GN000066; SPDB:SP  
A:Experimental source: strain 972h-; cosmid c56F8  
C:Genetics:  
A:Gene: SPDB:SPAC56F8.04C  
A:Map position: 1  
A:Genome: nuclear  
C:Keywords: mitochondrion; transferase

Query Match 80.0%; Score 32; DB 2; Length 358;  
Best Local Similarity 77.8%; Pred. No. 29;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
|||:|  
Db 311 RLSSMIYKV 319

## RESULT 8

A71277  
arginine-tRNA ligase (EC 6.1.1.19) (args) - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: A71277  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: A71277  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-589 <COL>  
A:Cross-references: UNIPROT:O83803; GB:AE001253; GB:AE000520; NID:G3323133; PIDN:NAC6579  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0831  
C:Superfamily: Bacillus arginine-tRNA ligase  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 80.0%; Score 32; DB 2; Length 589;  
Best Local Similarity 66.7%; Pred. No. 47;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
|:|:|:|  
Db 473 RLSSLUKKV 481

## RESULT 9

T50316  
hypothetical Armadillo/beta-catenin domain protein [imported] - fission yeast (Schizosac  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jul-2004  
C:Accession: T50316  
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Galibert, R  
submitted to the EMBL Data Library, January 2000  
A:Reference number: Z25061

A:Accession: T50316  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-664 <MCD>  
A:Cross-references: UNIPROT:Q9P7W7; EMBL:AL136536; PIDN:CAB66447.1; GSPDB:GN000067; SPDB:S  
A:Experimental source: strain 972h(-); cosmid c1703  
C:Genetics:  
A:Gene: SPDB:SPBC1703.03C  
A:Map position: 2  
A:Introns: 43/3

Query Match 80.0%; Score 32; DB 2; Length 664;  
Best Local Similarity 66.7%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
|||:|  
Db 624 RLKXNVKKI 632

## RESULT 10

H85470  
hypothetical protein AT4G39750 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: H85470  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: H85470  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-912 <STO>  
A:Cross-references: UNIPROT:Q9LDK6; GB:NC\_001268; NID:G7270959; PIDN:CAB80638.1; GSPDB:G  
C:Genetics:  
A:Gene: AT4G39750  
A:Map position: 4

Query Match 77.5%; Score 31; DB 2; Length 912;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
|||:|  
Db 403 RFSSMIRKI 411

## RESULT 11

T05012  
hypothetical protein T19P19.140 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05012  
R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hohenseil, J.; Mewes, H.W.;  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: Z15394  
A:Accession: T05012  
A:Molecule type: DNA  
A:Residues: 1-925 <BEV>  
A:Cross-references: UNIPROT:Q9LDK6; EMBL:AL022605  
A:Experimental source: cultivar Columbia; BAC clone T19P19  
C:Genetics:  
A:Map position: 4  
A:Introns: 142/3; 193/1; 551/1  
A:Note: T19P19.140

Query Match 77.5%; Score 31; DB 2; Length 925;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
|||:|

Db 416 RFSSMIRKI 424

## RESULT 12

T02473  
hypotheical protein At2g45740 [imported] - Arabidopsis thaliana  
N:Alternate names: hypotheical protein F4118.28  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T02473; C84894  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, August 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F4118 genomic sequence.  
A:Reference number: Z14674  
A:Accession: T02473  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-239 <ROU>  
A:Cross-references: UNIPROT:O80845; EMBL:AC004665; NID:g3386593; PID:g3386621  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84894  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-239 <STO>  
A:Cross-references: GB:AE002093; NID:g3386621; PIDN:AAC28551.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g45740; F4118.28  
A:Map position: 2  
A:introns: 65/3; 90/3; 115/3; 142/3; 167/3

Query Match 75.0%; Score 30; DB 2; Length 239;  
Best Local Similarity 66.7%; Pred. No. 56;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMWKKV 9

Db 148 RLSSSMKKI 156

## RESULT 13

T32363  
hypotheical protein C08E3.12 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32363  
R:Miller, N.; Kramer, J.; Keppler, D.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of C. elegans cosmid C08E3.  
A:Reference number: Z21155  
A:Accession: T32363  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-240 <MIL>  
A:Cross-references: UNIPROT:O17203; EMBL:AF025457; PIDN:AAB70970.1; GSPDB:GN00020; CESP:  
A:Experimental source: strain Bristol N2; clone C08E3  
C:Genetics:  
A:Gene: CESP:C08E3.12  
A:Map position: 2  
A:introns: 167/3

Query Match 75.0%; Score 30; DB 2; Length 240;  
Best Local Similarity 66.7%; Pred. No. 56;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMWKKV 9

Db 193 RIPSMLKKV 201

## RESULT 14

156519  
taipoxin-associated calcium binding protein-49 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I56519  
R:Dodds, D.; Schlimgen, A.K.; Lu, S.Y.; Perin, M.S.  
J. Neurochem. 64, 2339-2344, 1995  
A:Title: Novel reticular calcium binding protein is purified on taipoxin columns.  
A:Reference number: I56519; MUID:95239201; PMID:7722520  
A:Accession: I56519  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-318 <RES>  
A:Cross-references: UNIPROT:Q62703; EMBL:U15734; NID:G606967; PIDN:AAA80197.1; PID:G60696  
C:Superfamily: reticulocalbin; calmodulin repeat homology  
C:Keywords: calcium binding; EF hand  
F:62-94/Domain: calmodulin repeat homology <EF1>  
F:120-152/Domain: calmodulin repeat homology <EF2>  
F:187-219/Domain: calmodulin repeat homology <EF3>  
F:228-260/Domain: calmodulin repeat homology <EF4>  
F:264-296/Domain: calmodulin repeat homology <EF5>

Query Match 75.0%; Score 30; DB 2; Length 318;  
Best Local Similarity 55.6%; Pred. No. 73;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMWKKV 9

Db 66 RLQSIKKI 74

## RESULT 15

JC5402  
vitamin D receptor associated factor 1 - mouse  
N:Alternate names: ERC-55 protein  
C:Species: Mus musculus (house mouse)  
C:Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 09-Jul-2004  
C:Accession: JC5402  
R:Imai, T.; Matsuda, K.; Shimojima, T.; Hashimoto, T.; Masuhiro, Y.; Kitamoto, T.; Sugita  
Biochem. Biophys. Res. Commun. 233, 765-769, 1997  
A:Title: ERC-55, a binding protein for the papilloma virus E6 oncoprotein, specifically i  
A:Reference number: JC5402; MUID:97312489; PMID:9168930  
A:Accession: JC5402  
A:Molecule type: mRNA  
A:Residues: 1-319 <IMA>  
A:Cross-references: UNIPROT:O70341  
C:Comment: This protein acts as a vitamin D receptor-specific cofactor modulating its fur  
C:Superfamily: reticulocalbin; calmodulin repeat homology  
C:Keywords: EF hand  
F:63-95/Domain: calmodulin repeat homology <EF1>  
F:99-131/Domain: calmodulin repeat homology <EF2>  
F:151-183/Domain: calmodulin repeat homology <EF3>  
F:188-220/Domain: calmodulin repeat homology <EF4>  
F:229-261/Domain: calmodulin repeat homology <EF5>  
F:265-297/Domain: calmodulin repeat homology <EF6>

Query Match 75.0%; Score 30; DB 2; Length 319;  
Best Local Similarity 55.6%; Pred. No. 74;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLSSMWKKV 9

Db 67 RLQSIKKI 75

Search completed: July 13, 2005, 17:31:26  
Job time : 15.5556 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds

(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-32

Perfect score: 40

Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	181	2	Q8NFV7	Q8nf77 homo sapien
2	40	100.0	253	1	KLK7_HUMAN	P49862 homo sapien
3	40	100.0	253	2	Q8NSN9	Q8nsn9 homo sapien
4	33	82.5	202	2	Q6VPM5	Q6vpm5 anopheles g
5	33	82.5	273	2	Q7Q1V3	Q7q1v3 anopheles g
6	33	82.5	282	2	Q9S2Q8	Q9s2q8 streptomyce
7	33	82.5	297	2	Q7P141	Q7p141 chromobacte
8	33	82.5	372	1	COQ2_YEAST	P32378 saccharomyc
9	33	82.5	417	1	ORI_ANOGA	Q8wte7 anopheles g
10	33	82.5	530	2	Q6FR58	Q6fr58 candida gla
11	33	82.5	560	1	YM23_YEAST	P40215 saccharomyc
12	32	80.0	107	2	Q616B7	Q616b7 escherichia
13	32	80.0	169	2	Q899C6	Q899c6 clostridium
14	32	80.0	288	2	Q6AIB5	Q6aib5 desulfotale
15	32	80.0	330	2	Q7NVJ5	Q7nvj5 chromobacte
16	32	80.0	350	1	YUFN_BACSU	O05252 bacillus su
17	32	80.0	351	2	Q97Z26	Q97z26 sulfolobus
18	32	80.0	358	1	COQ2_SCHPO	O10252 schizosacch
19	32	80.0	589	1	SYR_TREPA	O83803 treponema p
20	32	80.0	664	2	Q9P7W7	Q9p7w7 schizosacch
21	31	77.5	97	2	Q9B021	Q9b021 bacterioph
22	31	77.5	355	2	Q67Q22	Q67q22 symbiobacte
23	31	77.5	390	2	Q8LEQ6	Q8leq6 arabidopsis
24	31	77.5	550	2	Q8W607	Q8w607 bacterioph
25	31	77.5	589	2	Q7XS95	Q7xs95 oryza sativ
26	31	77.5	610	2	Q72M04	Q72m04 leptospira
27	31	77.5	610	2	Q8EYH8	Q8eyh8 leptospira
28	31	77.5	751	2	Q7XMH8	Q7xmh8 oryza sativ
29	31	77.5	814	2	Q6L2C3	Q6l2c3 picophilus
30	31	77.5	912	2	Q9LDC6	Q9ldc6 arabidopsis
31	31	77.5	2159	2	Q8RVLL	Q8rvll zea mays (m

32	31	77.5	2159	2	Q8RVLL	Q8rvll zea mays (m
33	30	75.0	103	2	Q7JLJ1	Q7jil1 wolbachia p
34	30	75.0	105	2	Q3W83	Q3w83 arabidopsis
35	30	75.0	124	2	Q8BP39	Q8bp39 mus musculu
36	30	75.0	131	2	Q851X1	Q851x1 oryza sativ
37	30	75.0	207	2	Q6AV20	Q6av20 oryza sativ
38	30	75.0	217	2	Q73E11	Q73e11 bacillus ce
39	30	75.0	236	2	Q80845	Q80845 arabidopsis
40	30	75.0	240	2	Q17203	Q17203 caenorhabdi
41	30	75.0	308	2	Q7SK11	Q7sk11 dictyosteli
42	30	75.0	318	1	RCN2_RAT	Q62703 rattus norv
43	30	75.0	320	2	Q6P6X5	Q6p6x5 rattus norv
44	30	75.0	320	2	Q8BP92	Q8bp92 mus musculu
45	30	75.0	321	2	Q70341	Q70341 mus musculu

#### ALIGNMENTS

##### RESULT 1

Q8NFV7 PRELIMINARY; PRT; 181 AA.  
AC Q8NFV7;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Kallikrein 7 short variant protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A.  
RP TISSUE=Ovarian carcinoma;  
RX MEDLINE=22623266; PubMed=12738725;  
RA Dong Y., Kaulsh A., Brattsand M., Nicklin J., Clements J.A.;  
RT "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer  
produces novel variants with potential as cancer biomarkers.";  
RL Clin. Cancer Res. 9:1710-1720(2003).  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AF411215; AAN03663.1; -;  
DR HSSP; P00760; 1EZX.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 181 AA; 19887 MW; 86A28A03B80C2D78 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 181;

Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
| | | | | | | | | |

Db 51 RLSSMVKKV 59

##### RESULT 2

KLK7\_HUMAN STANDARD; PRT; 253 AA.  
ID KLK7\_HUMAN  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic  
 DN enzyme) (hSCCE).  
 OS Name=KLK7; Synonyms=PRSS6, SCCE;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
 RC TISSUE=Skin;  
 RX MEDLINE=94308225; PubMed=8034709;  
 RA Hanson L., Stromqvist M., Backman A., Wallbrandt P., Carlstein A.,  
 RA Egelrud T.;  
 RT "Cloning, expression, and characterization of stratum corneum  
 RT chymotryptic enzyme. A skin-specific human serine proteinase.";  
 RL J. Biol. Chem. 269:19420-19426 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Keratinocytes;  
 RA Yousef G.M., Scorilas A., Diamandis E.P.;  
 RT "Molecular characterization, mapping and tissue expression of the  
 RT human stratum corneum chymotryptic enzyme gene.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepker B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130 (2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Hanson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
 RA Wallbrandt P., Egelrud T.;  
 RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
 RT mice; a model for chronic itchy dermatitis.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95314630; PubMed=7794273;  
 RA Skyt A., Stromqvist M., Egelrud T.;  
 RT "Primary substrate specificity of recombinant human stratum corneum  
 RT chymotryptic enzyme.";  
 RL Biochem. Biophys. Res. Commun. 211:586-589 (1995).  
 CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive  
 CC structures in the cornified layer of the skin in the continuous  
 CC shedding of cells from the skin surface. Specific for amino acid  
 CC residues with aromatic side chains in the P1 position. SCCE  
 CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
 CC phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
 CC activation of precursors to inflammatory cytokines.  
 CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
 CC also seen in the brain and kidney.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
 CC subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; L33404; AAC37551.1; -;  
 DR EMBL; AF166330; AAD49718.1; -;  
 DR EMBL; AF243527; AAG33360.1; -;  
 DR EMBL; AF332583; AAK69624.1; -;  
 DR FIR; A53968; A53968.  
 DR HSP; P00760; IEX2.  
 DR MEROPS; S01.300; -;

DR Genew; HGNC:6368; KLK7.  
 DR H-InvDB; HIX0015373; -;  
 DR MIM; 604438; -;  
 DR GO; GO:0008236; P:serine-type peptidase activity; TAS.  
 DR GO; GO:0008544; P:epidermal differentiation; TAS.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
 KW Signal; Zymogen.  
 FT SIGNAL 1 22  
 FT PROPEP 23 29 Activation peptide.  
 FT CHAIN 30 253 Kallikrein 7.  
 FT ACT\_SITE 70 70 Charge relay system (By similarity).  
 FT ACT\_SITE 112 112 Charge relay system (By similarity).  
 FT ACT\_SITE 205 205 Charge relay system (By similarity).  
 FT DISULFID 36 137 By similarity.  
 FT DISULFID 55 71 By similarity.  
 FT DISULFID ? 239 By similarity.  
 FT DISULFID 144 211 By similarity.  
 FT DISULFID 176 190 By similarity.  
 FT DISULFID 201 226 By similarity.  
 FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 253 AA; 27525 MW; 268B8B15A76A668 CRC64;  
 Query Match 100.0%; Score 40; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RLSSMWKKV 9  
 DB 123 RLSSMWKKV 131  
 RESULT 3  
 Q8N5N9  
 ID Q8N5N9 PRELIMINARY; PRT; 253 AA.  
 AC Q8N5N9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Stratum corneum chymotryptic enzyme, preproprotein.  
 GN Name=KLK7;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.B., Schnerch A., Schein J.E.,

```
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RC SEQUENCE FROM N.A.
RP Strausberg R.;
RC TISSUE=Skin;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RL
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; BC032005; AAH32005.1; -.
DR HSSP; P00760; 1EXZ.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004293; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR Pfam; PF00089; Trypsin; 1_Ser_Cys.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Protease; Serine protease.
KW SEQUENCE 253 AA; 27608 MW; 2D68BGA41B22A668 CRC64;
SQ
Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131
|||||
RESULT 4
Q6VPM5 PRELIMINARY; PRT; 202 AA.
AC Q6VPM5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Odorant receptor 1 (Fragment).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4ARR, Yaounde, and L3-5;
RA Morlais I., Poncon N., Simard F., Cohuet A., Fontenille D.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY334006; AAR01131.1; -.
DR EMBL; AY334007; AAR01132.1; -.
DR EMBL; AY334008; AAR01130.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR004117; 7tm6_olfrrecept.
DR Pfam; PF02949; 7tm_6; 1.
KW Receptor.
FT NON_TER 1 1.
FT NON_TER 202 202
FT SEQUENCE 202 AA; 23639 MW; E4F6334A7FC8BEA0 CRC64;
SQ
Query Match 82.5%; Score 33; DB 2; Length 202;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 RLSSMVKKV 9
Db 185 RLSSMVKKV 193
|||||
```

```
RESULT 5
Q7QIV3 PRELIMINARY; PRT; 273 AA.
AC Q7QIV3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AgCP3543 (fragment).
GN Name=agCG53292; ORFNames=ENSANGG00000008607;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB0100807; EAA04012.1; -.
DR HSSP; P07359; 1GWB.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR_1; 5.
DR PRINTS; PR00019; LEUR1CHRP.
FT NON_TER 1 1
FT SEQUENCE 273 AA; 30363 MW; 413191DECB25D4C6 CRC64;
SQ
Query Match 82.5%; Score 33; DB 2; Length 273;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLSSMVKKV 9
Db 84 RVASVKKI 92
|||||
RESULT 6
Q9S2Q8 PRELIMINARY; PRT; 282 AA.
AC Q9S2Q8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 26, Last annotation update)
DE Putative endo alpha-1,4 polygalactosaminidase.
GN ORFNames=SC5P7.23c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RC "Complete genome sequence of the model actinomycete Streptomyces
RT Coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AL939111; CAB51262.1; -.
DR PIR; T35294; T35294.
KW Complete proteome.
SQ SEQUENCE 282 AA; 30789 MW; BEECFE1743703A33 CRC64;
```

```

Query Match      82.5%; Score 33; DB 2; Length 282;
Best Local Similarity 87.5%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLSSMVKK 8
DB      265 RLSSMLKK 272

RESULT 7
Q7P141
ID Q7P141 PRELIMINARY; PRT; 297 AA.
AC Q7P141;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CV0373;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C. de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Astaripe J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Faicao C.L.,
RA Fancinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangier T.B.,
RA Grattapaglia D., Grisead E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA Di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seidman E.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wasseem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability."
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AE016911; AAQ58051.1; -.
DR GO; GO:0045298; C:tubulin; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003924; F:GTPase activity; IEA.
DR GO; GO:0046785; F:microtubule polymerization; IEA.
DR InterPro; IPR002790; DUF88.
DR InterPro; IPR008280; Tub_FtsZ_C.
DR Pfam; PF01936; DUF88; 1.
DR ProDom; PD010896; DUF88; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 32462 MW; FCD2708DB087690 CRC64;

Query Match      82.5%; Score 33; DB 2; Length 297;
Best Local Similarity 77.8%; Pred. No. 92;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLSSMVKKV 9
DB      225 RLSSQWVKXL 233

RESULT 8
COQ2_YEAST STANDARD; PRT; 372 AA.
ID COQ2_YEAST Q6B113;
AC P32378; Q6B113;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Para-hydroxybenzoate-polyprenyltransferase, mitochondrial precursor
DE (EC 2.5.1.-) (PHB:polyprenyltransferase).
GN Name=COQ2; OrderedLocusNames=YNR041C; ORFNames=N3419;
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92156158; PubMed=1740455;
RA Ashby M.N., Kutsunai S.Y., Ackerman S., Tzagoloff A., Edwards P.A.;
RT "COQ2 is a candidate for the structural gene encoding para-
RT hydroxybenzoate:polyprenyltransferase.",
RL J. Biol. Chem. 267:4128-4136(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97313269; PubMed=9169873;
RA Philippsen P., Kleine K., Pohlmann R., Duesterhoeft A., Hamberg K.,
RA Hegemann J.H., Obermaier B., Urrestazu L.A., Aert R., Albermann K.,
RA Altmann R., Andre B., Baladron V., Ballesta J.P.G., Becam A.-M.,
RA Beinhaus J., Boskovic J., Buitrago M.J., Busseure F., Coster F.,
RA Crouzet M., D'Angelo M., Dal Pero F., De Antoni A., Del Rey F.,
RA Doignon F., Domdey H., Dubois E., Fiedler T., Fleig U., Floeth M.,
RA Fritz C., Gaillardin C., Garcia-Cantalejo J.M., Glansdorff N.,
RA Goffeau A., Guelidener U., Herbert C.J., Heumann K., Heuss-Neitzel D.,
RA Hilbert H., Hinni K., Iraqui Housaini I., Jaquet M., Jimenez A.,
RA Jonniaux J.-L., Karpfinger-Hartl L., Lanfranchi G., Lepingle A.,
RA Levesque H., Lyck R., Maftahi M., Mallet L., Maurer K.C.T.,
RA Messenguy F., Mewes H.-W., Moestl D., Nasr F., Nicaud J.-M.,
RA Niedenthal R.K., Pandolfo D., Pierard A., Piravandi E., Planta R.J.,
RA Pohl T.M., Purnelle B., Rebischung C., Remacha M., Revuelta J.L.,
RA Rinke M., Saiz J.E., Sartorello F., Scherens B., Sen-Gupta M.,
RA Soler-Mira A., Urbanus J.H.M., Valle G., Van Dyck L., Verhaesselt P.,
RA Viereindeels F., Vissers S., Voet M., Volckaert G., Wach A.,
RA Wambutt R., Wedler H., Zollner A., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIV
RT and its evolutionary implications.",
RL Nature 387:93-98(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,
RA Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley F.,
RA Zuo D., Raphael J., Hogle C., Jepeon D., Williamson J., Camargo A.,
RA Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
RA Labaer J.;
RT "Creation of the YFLEX clone resource: cloning of Saccharomyces
RT cerevisiae ORFs in the Gateway recombinational cloning system.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: Catalyzes the prenylation of para-hydroxybenzoate with
CC an all-trans polyprenyl group.
CC -!- PATHWAY: Coenzyme Q biosynthesis; second step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- SIMILARITY: Belongs to the ubiA prenyltransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```

DR EMBL; M81698; AAA34507.1; -.  
DR EMBL; 271656; CA936321.1; -.  
DR EMBL; AY693097; AAT93116.1; -.  
DR PIR; S20056; S20056.  
DR IntAct; P32378; -.  
DR GERMOnline; 143386; -.  
DR SGD; S00005324; COQ2.  
DR GO; GO:0004659; F:prenyltransferase activity; IMP.  
DR GO; GO:0006743; F:ubiquitinone metabolism; IMP.  
DR InterPro; IPR000537; UbA\_prenyltrans.  
DR InterPro; IPR006370; UbA\_proteo.  
DR Pfam; PF01040; UbA; 1.  
DR TIGRFam; TIGR01474; ubA\_proteo; 1.  
DR PROSITE; PS00943; UbA; 1.  
KW Isoprene biosynthesis; Mitochondrion; Transferase; Transit peptide;  
KW Transmembrane.  
FT TRANSIT 1 22 Mitochondrion (Potential).  
FT CHAIN 23 372 Para-hydroxybenzoate--  
FT polyprenyltransferase.  
FT TRANSMEM 92 112 1 (Potential).  
FT TRANSMEM 114 133 2 (Potential).  
FT TRANSMEM 170 191 3 (Potential).  
FT TRANSMEM 193 210 4 (Potential).  
FT TRANSMEM 229 249 5 (Potential).  
FT TRANSMEM 258 318 6 (Potential).  
FT DOMAIN 29 49 Ser-rich.  
FT DOMAIN 134 156 Alkyl polyprenyl diphosphate-binding  
FT site (Potential).  
FT CONFLICT 273 273 A -> T (in Ref. 3).  
FT SEQUENCE 372 AA; 41001 MW; 4D9738CE248B4AD6 CRC64;  
Query Match 82.5%; Score 33; DB 1; Length 372;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RLSSMVKKV 9  
Db 326 RLFSWIKKV 334  
RESULT 9  
ORL\_ANOGA STANDARD; PRT; 417 AA.  
AC Q8WTE7; Q7Q1T3;  
DT 29-MAR-2004 (Rel. 43, Created)  
DT 29-MAR-2004 (Rel. 43, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Odorant receptor Or1 (AgOri).  
GN Name=Or1;  
OS Anopheles gambiae (African malaria mosquito).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.  
OX NCBI\_TaxID=7165;  
RN [1]\_TaxID=7165;  
RP SEQUENCE FROM N.A., FUNCTION, INDUCTION, AND TISSUE SPECIFICITY.  
RC STRAIN=G3;  
RX MEDLINE=21593012; PubMed=11724964; DOI=10.1073/pnas.261432998;  
RA Fox A.N., Pitts R.J., Robertson H.M., Carlson J.R., Zwiebel L.J.;  
RT "Candidate odorant receptors from the malaria vector mosquito  
Anopheles gambiae and evidence of down-regulation in response to blood  
feeding.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:14693-14697(2001).  
RN [2]  
RC STRAIN=PEST;  
RP SEQUENCE FROM N.A.  
RG Anopheles genome sequencing consortium;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP FUNCTION, AND INDUCTION.  
RX PubMed=14724626; DOI=10.1038/427212a;  
RA Hallem E.A., Fox A.N., Zwiebel L.J., Carlson J.R.;  
RT "Olfaction: mosquito receptor for human-sweat odorant.";  
RL Nature 427:212-213(2004).

RN IDENTIFICATION, AND TISSUE SPECIFICITY.  
RX PubMed=12364795; DOI=10.1126/science.1076196;  
RA Hill C.A., Fox A.N., Pitts R.J., Kent L.B., Tan P.L., Chrystal M.A.,  
RA Cravchik A., Collins F.H., Robertson H.M., Zwiebel L.J.;  
RT "G protein-coupled receptors in Anopheles gambiae";  
RL Science 298:176-178(2002).  
CC -!- FUNCTION: Plays a critical role in the anthropophilic host-seeking  
behavior; establishes the host preference to transmit malaria. May  
participate in the phenomenon of decreased host-seeking behavior  
in disease vector mosquitoes after blood feeding.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Female-specific antennae and maxillary palp  
expression.  
CC -!- INDUCTION: Strong response to the odorant 4-methylphenol, a  
component of human sweat, when expressed in odorant receptor  
deficient Drosophila. In vivo, decreased expression in antennae  
after a blood meal.  
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor Dr-or  
family.  
CC -!- DATABASE: NAME=Protein Spotlight; NOTE=Issue 44 of February 2004;  
WWW="http://www.expasy.org/spotlight/articles/spt1044.html".  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
EMBL; AF364130; AAL35506.1; -.  
DR EMBL; AAB01008980; EAA13838.1; -.  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0005549; F:odorant binding; IDA.  
DR GO; GO:0004984; F:olfactory receptor activity; IDA.  
DR GO; GO:0042048; P:olfactory behavior; IDA.  
DR GO; GO:0007608; P:perception of smell; IDA.  
DR InterPro; IPR004117; 7tm6\_olfrrecept.  
DR InterPro; IPR008972; Cupredoxin.  
DR Pfam; PF02949; 7tm6; 1.  
DR G-protein coupled receptor; Multigene family; Olfaction; Receptor;  
KW Transmembrane.  
FT DOMAIN 1 2 Extracellular (Potential).  
FT TRANSMEM 3 23 1 (Potential).  
FT DOMAIN 24 45 Cytoplasmic (Potential).  
FT TRANSMEM 46 66 2 (Potential).  
FT DOMAIN 67 73 Extracellular (Potential).  
FT TRANSMEM 74 94 3 (Potential).  
FT DOMAIN 95 133 Cytoplasmic (Potential).  
FT TRANSMEM 134 154 4 (Potential).  
FT DOMAIN 155 178 Extracellular (Potential).  
FT TRANSMEM 179 199 5 (Potential).  
FT DOMAIN 200 284 Cytoplasmic (Potential).  
FT TRANSMEM 285 305 6 (Potential).  
FT DOMAIN 306 317 Extracellular (Potential).  
FT TRANSMEM 318 338 7 (Potential).  
FT DOMAIN 339 417 Cytoplasmic (Potential).  
SQ SEQUENCE 417 AA; 48520 MW; F60D7B7BD93D37F2 CRC64;  
Query Match 82.5%; Score 33; DB 1; Length 417;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RLSSMVKKV 9  
Db 219 RLGSWVKKL 227  
RESULT 10  
Q6FR58 PRELIMINARY; PRT; 530 AA.  
ID Q6FR58  
AC Q6FR58;

```
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P40215 Saccharomyces cerevisiae YMR145c NDH1.
GN ORFNames=CAGL0100748g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Lafontaine I., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barthe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykascen C.,
RA Boisrame A., Boyer E., Fairhead C., Confaniolieri F., de Daruvar A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zentou-Meyer M., Zivanovic I., Boitot-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380955; CAG60223.1; -.
DR GO; GO:0015036; P:disulfide oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001327; FAD_pyr_redox.
DR Pfam; PF00070; Pyr_redox.1
SQ SEQUENCE 530 AA; 5975 MW; 9BF95A03A0347248 CRC64;

Query Match 82.5%; Score 33; DB 2; Length 530;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9
DB 316 RLKTMVKV 324

RESULT 11
YM23 YEAST STANDARD; PRT; 560 AA.
AC P40215;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical 62.8 kDa protein in RPS16A-TIF34 intergenic region.
GN OrderedLocNames=YMR145C; ORFNames=YMR375.14C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII.";
RL Nature 387:90-93(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,

Hu Y., Vannberg F., Weger J., Kramer J., Moreira D., Kelley F.,
RA Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,
RA Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
RA LaBaer J.;
RT "Creation of the YFLEX clone resource: cloning of Saccharomycetes
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; Z47071; CAA87359.1; -.
CC EMBL; AY692785; AAT92804.1; -.
CC PIR; S50401; S50401.
CC GerMOnline; 142815; -.
CC SGD; S000004753; NDB1.
CC DR GO; GO:0005739; C:mitochondrion; IDA.
CC DR GO; GO:0003934; F:NADH dehydrogenase activity; IDA.
CC DR GO; GO:0019655; P:ethanol fermentation; IMP.
CC DR GO; GO:0006116; P:NADH oxidation; IDA.
CC DR InterPro; IPR001327; FAD_pyr_redox.
CC DR Pfam; PF00070; Pyr_redox; 1.
CC DR FAD; Flavoprotein; Hypothetical protein; NAD; Oxidoreductase;
KW Ubiquinone.
SQ SEQUENCE 560 AA; 62774 MW; 10B1795E1E29C34 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 560;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9
DB 347 RLKTMVKV 355

RESULT 12
Q616B7
ID Q616B7 PRELIMINARY; PRT; 107 AA.
AC Q616B7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Escherichia coli.
OG Plasmid R388.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Anbutsu H., Tsuda M.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB182370; BAD24124.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 107 AA; 12021 MW; 1B128007B7367B5B CRC64;

Query Match 80.0%; Score 32; DB 2; Length 107;
Best Local Similarity 77.8%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9
DB 52 RVSSMWKSV 60

RESULT 13
Q899C6 PRELIMINARY; PRT; 169 AA.
ID Q899C6
```



```

AC Q895C6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Anaerobic ribonucleoside-triphosphate reductase activating protein (EC
DE 1.97.1.1).
GN OrderedLocusNames=CTC00258;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;
RA Brueggemann H., Baerumer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease."
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015936; AAC03490.1; -.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006800; P:oxygen and reactive oxygen species metabolism; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001989; Radical activat.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; 1.
DR PROSITE; PS01190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS01087; RADICAL ACTIVATING; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 169 AA; 19292 MW; CCTAB9511D3844C1 CRC64;

Query Match 80.0%; Score 32; DB 2; Length 169;
Best Local Similarity 75.0%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVVK 8
Db 85 RLASMIKK 92

RESULT 14
QGAIB5 PRELIMINARY; PRT; 288 AA.
AC Q6AIB5;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Probable UTP-glucose-1-phosphate uridylyltransferase.
GN OrderedLocusNames=DPPB68;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSv54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Prickey T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments."
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522871; CAG37932.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0003983; F:UTP-glucose-1-phosphate uridylyltransferase. . .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0006011; P:UDP-glucose metabolism; IEA.
DR InterPro; IPR005771; Galu_trans.

Q895C6;
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
DR TIGRFAMs; TIGR01099; galU; 1.
KW Complete proteome; Nucleoside diphosphate transferase; Transferase.
SQ SEQUENCE 288 AA; 31782 MW; FID28AC4A04BE3CB CRC64;

Query Match 80.0%; Score 32; DB 2; Length 288;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVVK 8
Db 184 RLSSMVVK 191

RESULT 15
Q7NVJ5 PRELIMINARY; PRT; 330 AA.
ID Q7NVJ5;
AC Q7NVJ5;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DE Probable membrane lipoprotein.
GN OrderedLocusNames=CV2347;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Astaripe J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogio M., Bonatto S.,
RA Bordignon J., Brigo M.M., Brito C.A., Brocchi M., Burley H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Garzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wassen R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability."
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AE016918; AAQ60019.1; -.
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 330 AA; 35382 MW; 2APAB050961AF87F CRC64;

Query Match 80.0%; Score 32; DB 2; Length 330;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSMVVKV 9
Db 1:|||||

```

Db 247 LTSMVKKV 254

Search completed: July 13, 2005, 17:29:18  
Job time : 65.2222 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19,444 Seconds  
(without alignments)  
34,552 Million cell updates/sec

Title: US-09-905-083A-32

Perfect score: 40

Sequence: 1 RLSSWVKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	3	US-09-502-600-32
2	40	100.0	9	4	US-09-918-243-32
3	40	100.0	144	4	US-09-618-259-4
4	40	100.0	154	3	US-09-261-416-7
5	40	100.0	161	3	US-09-261-416-8
6	40	100.0	224	3	US-08-944-483-33
7	40	100.0	225	2	US-08-557-146-12
8	40	100.0	225	2	US-09-027-337-4
9	40	100.0	225	2	US-09-154-344-12
10	40	100.0	225	4	US-09-644-600-4
11	40	100.0	225	4	US-09-654-600A-4
12	40	100.0	253	2	US-08-557-146-2
13	40	100.0	253	2	US-08-824-874-3
14	40	100.0	253	2	US-09-154-344-2
15	40	100.0	253	3	US-08-930-188-2
16	40	100.0	253	3	US-09-210-084-3
17	40	100.0	253	4	US-09-764-762-3
18	40	100.0	253	5	PCT-US96-04294-2
19	40	100.0	265	4	US-09-949-016-7716
20	36	90.0	9	3	US-09-502-600-110
21	36	90.0	9	4	US-09-918-243-110
22	31	77.5	9	3	US-09-502-600-79
23	31	77.5	9	3	US-09-502-600-87
24	31	77.5	9	4	US-09-918-243-79
25	31	77.5	9	4	US-09-918-243-87
26	30	75.0	156	4	US-09-328-352-7506
27	29	72.5	571	4	US-09-690-942-15

Sequence 2, Appli  
Sequence 4, Appli  
Sequence 6, Appli  
Sequence 2383, A  
Sequence 7990, Ap  
Sequence 18044, A  
Sequence 3154, Ap  
Sequence 4867, Ap  
Sequence 49085, A  
Sequence 4737, Ap  
Sequence 6154, Ap  
Sequence 3650, Ap  
Sequence 1360, Ap  
Sequence 5465, Ap  
Sequence 2, Appli  
Sequence 2462, Ap  
Sequence 4562, Ap  
Patent No. 5171685

28 29 72.5 582 1 US-08-261-086-2  
29 29 72.5 582 1 US-08-261-086-4  
30 29 72.5 582 1 US-08-261-086-6  
31 28 70.0 90 4 US-09-248-796A-23383  
32 28 70.0 154 4 US-09-328-352-7990  
33 28 70.0 174 4 US-09-248-796A-18044  
34 28 70.0 214 4 US-09-583-110-3154  
35 28 70.0 217 4 US-09-107-433-4867  
36 28 70.0 218 4 US-09-270-767-49085  
37 28 70.0 250 4 US-09-107-532A-4737  
38 28 70.0 289 4 US-09-107-532A-6154  
39 28 70.0 341 3 US-09-134-001C-3650  
40 28 70.0 357 4 US-09-710-279-1360  
41 28 70.0 405 4 US-09-134-000C-5465  
42 28 70.0 459 3 US-09-491-785-2  
43 28 70.0 459 4 US-09-710-279-2462  
44 28 70.0 464 3 US-09-134-001C-4562  
45 28 70.0 565 6 5171685-6

## ALIGNMENTS

### RESULT 1

US-09-502-600-32  
; Sequence 32, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 32  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 123-131 of the SCCE protein

Query Match Similarity 100.0%; Score 40; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSWVKV 9  
Db 1 RLSSWVKV 9

### RESULT 2

US-09-918-243-32  
; Sequence 32, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 32  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens

;  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 123-131 of the SCCE protein  
US-09-918-243-32

Query Match 100.0%; Score 40; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 1 RLSSMVKKV 9

RESULT 3  
US-09-618-259-4  
; Sequence 4, Application US/09618259  
; Patent No. 6642013  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease  
; FILE REFERENCE: D6020CIP2  
; CURRENT APPLICATION NUMBER: US/09/618,259  
; CURRENT FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: US 09/127,444  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 72  
; SEQ ID NO 4  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: unknown  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic  
; OTHER INFORMATION: enzyme (scce) catalytic domain  
US-09-618-259-4

Query Match 100.0%; Score 40; DB 4; Length 144;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 60 RLSSMVKKV 68

RESULT 4  
US-09-261-416-7  
; Sequence 7, Application US/09261416A  
; Patent No. 6291663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease  
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease  
; FILE REFERENCE: D6192  
; CURRENT APPLICATION NUMBER: US/09/261,416A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 14  
; SEQ ID NO 7  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum  
; OTHER INFORMATION: chymotryptic enzyme (Scce) homologous to similar  
; OTHER INFORMATION: domain in TAGD-12  
US-09-261-416-7

Query Match 100.0%; Score 40; DB 3; Length 154;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 69 RLSSMVKKV 77

RESULT 5  
US-09-261-416-8  
; Sequence 8, Application US/09261416A  
; Patent No. 6291663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease  
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease  
; FILE REFERENCE: D6192  
; CURRENT APPLICATION NUMBER: US/09/261,416A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 14  
; SEQ ID NO 8  
; LENGTH: 161  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of hepsin (Heps)  
; OTHER INFORMATION: homologous to similar domain in TAGD-12  
US-09-261-416-8

Query Match 100.0%; Score 40; DB 3; Length 161;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 75 RLSSMVKKV 83

RESULT 6  
US-08-944-483-33  
; Sequence 33, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; TITLE OF INVENTION: OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:

Query Match 100.0%; Score 40; DB 3; Length 161;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6183.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 224 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6232456e  
US-08-944-483-33

Query Match 100.0%; Score 40; DB 3; Length 224;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 94 RLSSMVKKV 102  
|||||

RESULT 7  
US-08-557-146-12  
; Sequence 12, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION NUMBER: US/08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steiner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 225 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
US-08-557-146-12

Query Match 100.0%; Score 40; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 94 RLSSMVKKV 102  
|||||

RESULT 8  
US-09-027-337-4  
; Sequence 4, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hiroto  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in  
; TITLE OF INVENTION: Breast and Ovarian Carcinomas  
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027,337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to  
; OTHER INFORMATION: similar domain in TADG-15  
US-09-027-337-4

Query Match 100.0%; Score 40; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 95 RLSSMVKKV 103  
|||||

RESULT 9  
US-09-154-344-12  
; Sequence 12, Application US/09154344  
; Patent No. 5981256  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/154,344  
; FILING DATE: 16-SEP-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steiner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 225 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
US-09-154-344-12

Query Match 100.0%; Score 40; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 95 RLSSMVKKV 103  
|||||

ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6183.US.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 224 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6232456e  
US-08-944-483-33

Query Match 100.0%; Score 40; DB 3; Length 224;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 94 RLSSMVKKV 102  
|||||

RESULT 7  
US-08-557-146-12  
; Sequence 12, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION NUMBER: US/08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steiner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 225 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
US-08-557-146-12

Query Match 100.0%; Score 40; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 94 RLSSMVKKV 102  
|||||

RESULT 8  
US-09-027-337-4  
; Sequence 4, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hiroto  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in  
; TITLE OF INVENTION: Breast and Ovarian Carcinomas  
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027,337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to  
; OTHER INFORMATION: similar domain in TADG-15  
US-09-027-337-4

Query Match 100.0%; Score 40; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 95 RLSSMVKKV 103  
|||||

RESULT 9  
US-09-154-344-12  
; Sequence 12, Application US/09154344  
; Patent No. 5981256  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/154,344  
; FILING DATE: 16-SEP-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steiner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 225 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
US-09-154-344-12

Query Match 100.0%; Score 40; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RLSSMVKKV 9  
|||||

Db 95 RLSSMVKKV 103  
|||||

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-154-344-12

Query Match          100.0%; Score 40; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 10
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-644-600-4

Query Match          100.0%; Score 40; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 11
US-09-654-600A-4
; Sequence 4, Application US/09654600A
; Patent No. 6649741
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/654,600A
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-654-600A-4

Query Match          100.0%; Score 40; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 12
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2

Query Match          100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 13
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```

; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
; US-08-824-874-3

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 15
US-08-930-188-2
; Sequence 2, Application US/08930188
; Patent No. 6093397
; GENERAL INFORMATION:
; APPLICANT: Dixon, Eric P.
; APPLICANT: Johnstone, Edward M.
; APPLICANT: Little, Shella P.
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
; TITLE OF INVENTION: RELATED NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,188
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/416,257
; FILING DATE: 04-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bialock, Donna K.
; REGISTRATION NUMBER: 38,082
; REFERENCE/DOCKET NUMBER: X9239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-277-1090
; TELEFAX: 317-276-3861
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-930-188-2

Query Match 100.0%; Score 40; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;

; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
; US-08-824-874-3

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 14
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
```

	Matches	9;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	RLSSMKKV	9							
Db	123	RLSSMKKV	131							

Search completed: July 13, 2005, 17:34:25  
Job time : 20.4444 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds

(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-32

Perfect score: 40

Sequence: 1 RLSSWVKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	40	100.0	9	9 US-09-918-243-32	Sequence 32, Appl
2	40	100.0	9	9 US-09-905-083-32	Sequence 32, Appl
3	40	100.0	9	15 US-10-372-521-32	Sequence 32, Appl
4	40	100.0	9	16 US-10-831-075-32	Sequence 32, Appl
5	40	100.0	30	16 US-10-831-075-137	Sequence 4, Appl
6	40	100.0	144	9 US-09-796-294-4	Sequence 4, Appl
7	40	100.0	144	14 US-10-461-787-4	Sequence 4, Appl
8	40	100.0	224	11 US-09-789-210-33	Sequence 33, Appl
9	40	100.0	224	15 US-10-262-511-104	Sequence 104, App
10	40	100.0	225	15 US-10-600-187-4	Sequence 4, Appl
11	40	100.0	247	15 US-10-262-511-102	Sequence 102, App

12	40	100.0	250	15	US-10-262-511-92	Sequence 92, Appl
13	40	100.0	252	15	US-10-262-511-94	Sequence 94, Appl
14	40	100.0	253	9	US-09-888-615-98	Sequence 98, Appl
15	40	100.0	253	9	US-09-764-762-3	Sequence 3, Appl
16	40	100.0	253	14	US-10-071-214-2	Sequence 2, Appl
17	40	100.0	253	14	US-10-071-214-48	Sequence 48, Appl
18	40	100.0	253	14	US-10-264-283-90	Sequence 90, Appl
19	40	100.0	253	15	US-10-295-027-498	Sequence 48, Appl
20	40	100.0	253	15	US-10-173-999-48	Sequence 48, Appl
21	40	100.0	253	16	US-10-408-765A-639	Sequence 639, App
22	40	100.0	253	16	US-10-643-795A-95	Sequence 95, Appl
23	40	100.0	253	17	US-10-948-518-95	Sequence 95, Appl
24	40	100.0	253	17	US-10-868-490A-1	Sequence 1, Appl
25	40	100.0	257	15	US-10-344-394-38	Sequence 38, Appl
26	36	90.0	9	9	US-09-918-243-110	Sequence 110, App
27	36	90.0	9	9	US-09-905-083-110	Sequence 110, App
28	36	90.0	9	15	US-10-372-521-110	Sequence 110, App
29	36	90.0	9	16	US-10-831-075-110	Sequence 110, App
30	33	82.5	84	15	US-10-424-599-244024	Sequence 244024,
31	33	82.5	369	16	US-10-437-963-155539	Sequence 155539,
32	33	82.5	394	14	US-10-094-240-4	Sequence 4, Appl
33	33	82.5	394	14	US-10-056-405-4	Sequence 4, Appl
34	32	80.0	258	15	US-10-425-114-61611	Sequence 61611, A
35	32	80.0	293	15	US-10-425-114-49273	Sequence 49273, A
36	32	80.0	295	15	US-10-425-114-43367	Sequence 43367, A
37	32	80.0	295	15	US-10-425-114-59862	Sequence 59862, A
38	32	80.0	296	15	US-10-425-114-53750	Sequence 53750, A
39	32	80.0	298	15	US-10-425-114-63676	Sequence 63676, A
40	32	80.0	589	15	US-10-282-122A-76609	Sequence 76609, A
41	31	77.5	9	9	US-09-918-243-79	Sequence 79, Appl
42	31	77.5	9	9	US-09-918-243-87	Sequence 87, Appl
43	31	77.5	9	9	US-09-905-083-79	Sequence 79, Appl
44	31	77.5	9	9	US-09-905-083-87	Sequence 87, Appl
45	31	77.5	9	15	US-10-372-521-79	Sequence 79, Appl

#### ALIGNMENTS

RESULT 1  
US-09-918-243-32  
; Sequence 32, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 32  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 123-131 of the SCCE protein  
US-09-918-243-32

Query Match 100.0%; Score 40; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RLSSWVKV 9  
Db 1 RLSSWVKV 9

RESULT 2

US-09-905-083-32  
; Sequence 32, Application US/09905083  
; Patent No. US20020146708A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; TITLE OF INVENTION: Ovarian Cancer  
; FILE REFERENCE: D6223CIP/C/Div  
; CURRENT APPLICATION NUMBER: US/09/905,083  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 09/502,600  
; PRIOR FILING DATE: 2000-02-11  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 32  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 123-131 of the SCCE protein  
US-09-905-083-32

Query Match 100.0%; Score 40; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 1 RLSSMVKKV 9

RESULT 3  
US-10-372-521-32  
; Sequence 32, Application US/10372521  
; Publication No. US20030223973A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP2  
; CURRENT APPLICATION NUMBER: US/10/372,521  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 32  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 123-131 of the SCCE protein  
US-10-372-521-32

Query Match 100.0%; Score 40; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 1 RLSSMVKKV 9

RESULT 4  
US-10-831-075-32  
; Sequence 32, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 32  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 123-131 of the SCCE protein  
US-10-831-075-32

Query Match 100.0%; Score 40; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 1 RLSSMVKKV 9

RESULT 5  
US-10-831-075-137  
; Sequence 137, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 137  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 110-139 of the SCCE protein  
US-10-831-075-137

Query Match 100.0%; Score 40; DB 16; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9  
Db 14 RLSSMVKKV 22

RESULT 6  
US-09-796-294-4  
; Sequence 4, Application US/09796294  
; Patent No. US20020037581A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Extracellular Serine Protease  
; FILE REFERENCE: D6020CIP3  
; CURRENT APPLICATION NUMBER: US/09/796,294  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 09/618,259  
; PRIOR FILING DATE: 2000-07-18  
; NUMBER OF SEQ ID NOS: 72  
; SEQ ID NO 4  
; LENGTH: 144

```
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-09-796-294-4

Query Match      100.0%; Score 40; DB 9; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 60 RLSSMVKKV 68

RESULT 7
US-10-461-787-4
; Sequence 4, Application US/10461787
; Publication No. US20030199010A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: No. US20030199010A1el Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/10/461,787
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US/09/618,259
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-10-461-787-4

Query Match      100.0%; Score 40; DB 14; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 60 RLSSMVKKV 68

RESULT 8
US-09-789-210-33
; Sequence 33, Application US/09789210
; Publication No. US20040241646A1
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GRANADOS, EDWARD N.
; KLASS, MICHAEL R.
; RUSSELL, JOHN C.
; STEWART, KENT D.
; STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
```

```
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/789,210
; FILING DATE: 20-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/944,483
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-789-210-33

Query Match      100.0%; Score 40; DB 11; Length 224;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 94 RLSSMVKKV 102

RESULT 9
US-10-262-511-104
; Sequence 104, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
```

```
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 104
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-104

Query Match      100.0%; Score 40; DB 15; Length 224;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 109 RLSSMVKKV 117

RESULT 10
US-10-600-187-4
; Sequence 4, Application US/10600187
; Publication No. US20040086910A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotoishi
; TITLE OF INVENTION: TAGG-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/10/600,187
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US/09/654,600A
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-10-600-187-4

Query Match      100.0%; Score 40; DB 15; Length 225;
```

```
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 11
US-10-262-511-102
; Sequence 102, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 102
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-262-511-102

Query Match 100.0%; Score 40; DB 15; Length 247;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 109 RLSSMVKKV 117

RESULT 12

US-10-262-511-92

; Sequence 92, Application US/10262511  
; Publication No. US20040038223A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerkhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-462C

; CURRENT APPLICATION NUMBER: US/10/262,511

; CURRENT FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/373,815

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,642

; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-03-16

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/373,260

; PRIOR FILING DATE: 2002-04-17

; PRIOR APPLICATION NUMBER: 60/373,826

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,435

; PRIOR FILING DATE: 2001-10-05

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 439

; SOFTWARE: CuraseqList version 0.1

; SEQ ID NO 92

; LENGTH: 250

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-262-511-92

Query Match 100.0%; Score 40; DB 15; Length 250;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 120 RLSSMVKKV 128

RESULT 13

US-10-262-511-94

; Sequence 94, Application US/10262511

; Publication No. US20040038223A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerkhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-462C

; CURRENT APPLICATION NUMBER: US/10/262,511

; CURRENT FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/373,815

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,642

; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/373,260

; PRIOR FILING DATE: 2002-04-17

; PRIOR APPLICATION NUMBER: 60/373,826

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,435

; PRIOR FILING DATE: 2001-10-05

; Remaining Prior Application data removed - See File Wrapper or PALM.

```
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 94
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-94

Query Match      100.0%; Score 40; DB 15; Length 252;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 122 RLSSMVKKV 130

RESULT 14
US-09-888-615-98
; Sequence 98, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Query Match      100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 15
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-764-762-3

Query Match      100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

Search completed: July 13, 2005, 18:23:33
Job time : 69.8889 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083A-33

Perfect score: 40

Sequence: 1 LLLPLQLL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 16Dec04: \*  
1: Geneseqp1980s: \*  
2: Geneseqp1990s: \*  
3: Geneseqp2000s: \*  
4: Geneseqp2001s: \*  
5: Geneseqp2002s: \*  
6: Geneseqp2003as: \*  
7: Geneseqp2003bs: \*  
8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	AAE08238	Human str
2	40	100.0	9	ADR68794	Human str
3	40	100.0	136	ABG23378	Novel hum
4	40	100.0	198	ADA05736	Human NOV
5	40	100.0	198	ADN62900	Human NOV
6	40	100.0	250	ADA05732	Human NOV
7	40	100.0	250	ADN62896	Human NOV
8	40	100.0	253	AAE67888	Human str
9	40	100.0	253	AAW05383	Human any
10	40	100.0	253	ABB84421	Human SCC
11	40	100.0	253	ABB84406	Human SCC
12	40	100.0	253	AAU82740	Amino aci
13	40	100.0	253	ABU07440	Protein d
14	40	100.0	253	ABU07471	Protein d
15	40	100.0	253	ABR58471	Human str
16	40	100.0	253	ADN80484	Ovarian c
17	40	100.0	253	ADJ68833	Human hea
18	40	100.0	253	ADN319180	Cancer/an
19	40	100.0	253	ADL06515	Human tum
20	40	100.0	253	ADN04182	Antipsori
21	40	100.0	253	ADR72880	Human ova
22	40	100.0	257	ABR21326	Human HSC
23	38	95.0	812	ABO65558	Klebsiell
24	36	90.0	9	AAE08240	Human str
25	36	90.0	9	AAE08241	Human str

26	36	90.0	9	ADR68796	Human str
27	36	90.0	9	ADR68797	Human str
28	35	87.5	61	AAO12472	Human pol
29	35	87.5	142	AAE63580	Human gas
30	35	87.5	156	AAE63578	Human gas
31	35	87.5	159	AAE63582	Human gas
32	34	85.0	23	ABB43858	Peptide #
33	34	85.0	23	AAE37771	Peptide #
34	34	85.0	23	AAE64837	Human bra
35	34	85.0	23	ABG59233	Human liv
36	34	85.0	23	ABG59233	Human liv
37	34	85.0	115	ABF59165	Human pol
38	34	85.0	201	ABB78636	Rat OST10
39	34	85.0	369	ADL91113	Bovine lu
40	34	85.0	369	ADL91117	Bovine lu
41	34	85.0	369	ADL91127	Bovine lu
42	34	85.0	371	AAE75642	Bovine co
43	34	85.0	371	ADL91107	Bovine lu
44	34	85.0	371	ADL91076	Bovine co
45	34	85.0	371	ADL91059	Bovine co

## ALIGNMENTS

## RESULT 1

AAE08238  
ID AAE08238 standard; peptide; 9 AA.

XX AAE08238;

AC AC

DT 01-NOV-2001 (first entry)

XX Human stratum corneum chymotrypsin enzyme peptide #3 (residues 5-13).

DE Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;  
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;  
KW antisense therapy; malignant hyperplasia.

XX Homo sapiens.

OS WO200159158-A1.

XX 16-AUG-2001.

PD 07-FEB-2001; 2001WO-US003977.

PF 11-FEB-2000; 2000US-00502600.

XX (UYAR-) UNIV ARKANSAS.

XX O'brien TJ;

XX WPI; 2001-514676/56.

XX Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX Claim 25; Page 102; 127pp; English.

XX The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

XX Sequence 9 AA;

SQ Query Match 100.0%; Score 40; DB 4; Length 9;

```
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
Db 1 LLLPLQILL 9

RESULT 2
ADR68794
ID ADR68794 standard; peptide; 9 AA.
XX AC
XX ADR68794;
XX DT
XX 02-DEC-2004 (first entry)
XX DE
XX Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:33.
XX KW
XX serine protease; stratum corneum chymotrytic enzyme; SCCE;
XX immune response; ovarian cancer; lung cancer; prostate cancer;
XX pancreatic cancer; colon cancer.
XX OS
XX Homo sapiens.
XX PN
XX WO2004075723-A2.
XX PD
XX 10-SEP-2004.
XX PF
XX 20-FEB-2004; 2004WO-US005134.
XX PR
XX 21-FEB-2003; 2003US-00372521.
XX PA
XX (UYAR-) UNIV ARKANSAS.
XX PI
XX O'brien TJ, Cannon MJ, Santin A;
XX DR
XX WPI; 2004-653294/63.
XX PT
XX Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for
XX vaccinating an individual against SCCE, and in monitoring and developing
XX immunotherapies for ovarian and other malignancies.
XX PS
XX Claim 5; SEQ ID NO 33; 117pp; English.
XX CC
XX The invention relates to a novel method for vaccinating an individual
XX against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating
XX an individual with a SCCE peptide, which elicits an immune response in
XX the individual. A peptide of the invention acts as a stratum corneum
XX chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating
XX an individual against SCCE, particularly an individual having, suspected
XX or at risk of getting ovarian, lung, prostate, pancreatic or colon
XX cancer. The oligonucleotide is useful for treating a neoplastic state in
XX an individual, such as ovarian, breast, lung, colon, prostate, or
XX pancreatic cancer, and other cancers in which SCCE is overexpressed. The
XX peptides are also useful in the monitoring and development of
XX immunotherapies for ovarian and other malignancies. The present sequence
XX represents a peptide fragment of serine protease SCCE (stratum corneum
XX chymotrytic enzyme).
XX SQ
XX Sequence 9 AA;
XX Query Match 100.0%; Score 40; DB 8; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
Db 1 LLLPLQILL 9

RESULT 3
ABG23378
ID ABG23378 standard; protein; 136 AA.
XX AC
```

---

```
XX AC
XX ABG23378;
XX DT
XX 18-FEB-2002 (first entry)
XX DE
XX Novel human diagnostic protein #23369.
XX KW
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS
XX Homo sapiens.
XX PN
XX WO200175067-A2.
XX PD
XX 11-OCT-2001.
XX PF
XX 30-MAR-2001; 2001WO-US008631.
XX PR
XX 31-MAR-2000; 2000US-00540217.
XX PR
XX 23-AUG-2000; 2000US-00649167.
XX PA
XX (HYSE-) HYSEQ INC.
XX PI
XX Drmanac RT, Liu C, Tang YT;
XX DR
XX WPI; 2001-639362/73.
XX DR
XX N-PSDB; AAS87565.
XX PT
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX PS
XX Claim 20; SEQ ID NO 53737; 103pp; English.
XX CC
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 136 AA;
XX Query Match 100.0%; Score 40; DB 4; Length 136;
XX Best Local Similarity 100.0%; Pred. No. 12;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
Db 11 LLLPLQILL 19

RESULT 4
ADA05736
ID ADA05736 standard; protein; 198 AA.
XX AC
XX ADA05736;
```





PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJJI/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERR/) ZERRHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG W.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGE E M L.  
 PA (BERG/) BERGHS C.  
 PA (DIP/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.  
 XX  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytke KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerrhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 WPI; 2004-213931/20.  
 DR N-PSDB; ADN62899.  
 DR  
 XX  
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 XX  
 PS Claim 1; SEQ ID NO 96; 395bp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX  
 SQ Sequence 198 AA;  
 Query Match 100.0%; Score 40; DB 8; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLLPLQLLL 9  
 DB 5 LLLPLQLLL 13  
 |||||  
 |||||  
 RESULT 6  
 ADA05732  
 ID ADA05732 standard; protein; 250 AA.  
 XX  
 AC ADA05732;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV18a protein SEQ ID NO:92.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytosolic; neurotropic; neuroprotective;  
 KW antiparkinsonian; antilipaeimic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003029424-A2.  
 PD 10-APR-2003.  
 XX  
 PF 02-OCT-2002; 2002WO-US031373.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327497P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.

12-OCT-2001; 2001US-0328849P.  
 15-OCT-2001; 2001US-0329414P.  
 17-OCT-2001; 2001US-0330142P.  
 18-OCT-2001; 2001US-0330309P.  
 22-OCT-2001; 2001US-0341058P.  
 24-OCT-2001; 2001US-0343629P.  
 29-OCT-2001; 2001US-0349575P.  
 01-NOV-2001; 2001US-0346357P.  
 17-APR-2002; 2002US-0373260P.  
 19-APR-2002; 2002US-0373815P.  
 19-APR-2002; 2002US-0373817P.  
 19-APR-2002; 2002US-0373826P.  
 19-APR-2002; 2002US-0373884P.  
 22-APR-2002; 2002US-0374977P.  
 16-MAY-2002; 2002US-0381037P.  
 16-MAY-2002; 2002US-0381038P.  
 17-MAY-2002; 2002US-0381042P.  
 28-MAY-2002; 2002US-0381642P.  
 29-MAY-2002; 2002US-0383656P.  
 25-JUN-2002; 2002US-0383831P.  
 01-OCT-2002; 2002US-0391335P.  
 (CURA-) CURAGEN CORP.  
 Smithson G, Millett I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Pattarajan M, Szyteck KA, Edinger SR, Ellerman K, Malyankar UM;  
 Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 Shimketa RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 WPI; 2003-381626/36.  
 N-PSDB; ADA05731.  
 New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 pharmacogenomics.  
 Claim 1; Page 169-170; 586pp; English.  
 The present invention describes NOVX proteins, where X can be 1 to 55  
 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 described above and a carrier; (2) a kit comprising, in one or more  
 containers, the composition described above; (3) an isolated nucleic acid  
 molecule which encodes a NOVX protein of the invention; (4) a vector  
 comprising the nucleic acid molecule described above; (5) a cell  
 comprising the above vector; (6) an antibody that immunospecifically  
 binds to the polypeptide described above; (7) methods for determining the  
 presence or amount of the above polypeptide or nucleic acid molecule in a  
 sample; (8) methods for determining the presence of or predisposition to  
 a disease associated with altered levels of expression of the above  
 polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 method of identifying an agent that binds to the polypeptide described  
 above; (10) a method for identifying a potential therapeutic agent for  
 use in treating a pathology that is related to an aberrant expression or  
 aberrant physiological interactions of the polypeptide; (11) a method of  
 screening for a modulator of activity or of latency or predisposition to  
 a pathology associated with the polypeptide; (12) a method for modulating  
 the activity of the polypeptide described above; (13) methods of treating  
 or preventing a pathology associated with the above polypeptide in a  
 mammal; and (14) a method for producing the above polypeptide. NOVX  
 sequences have antidiabetic, anorectic, antibacterial, virucide,  
 immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 and antilipemic activities, and can be used in gene therapy. The  
 polypeptide is useful in manufacturing a medicament for treating a  
 syndrome associated with a human disease. The polypeptide or the nucleic  
 acid molecule may be used to diagnose, treat or prevent metabolic  
 disorders such as diabetes or obesity, infections, cachexia, cancer,  
 neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 probes, in chromosome mapping, tissue typing, preventive medicine and  
 pharmacogenomics. The present sequence represents a human NOVX from the  
 present invention.  
 XX Sequence 250 AA;  
 Query Match 100.0%; Score 40; DB 6; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLLPLQLLL 9  
 Db 2 LLLPLQLLL 10  
 |||||  
 |||||  
 RESULT 7  
 ADM62896  
 ID ADM62896 standard; protein; 250 AA.  
 XX  
 AC ADM62896;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human NOV18a.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 Alzheimer's disease; Parkinson's disease; immune disorder;  
 haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 02-OCT-2001; 2001US-0326483P.  
 05-OCT-2001; 2001US-0327435P.  
 05-OCT-2001; 2001US-0327449P.  
 09-OCT-2001; 2001US-0327917P.  
 09-OCT-2001; 2001US-0328029P.  
 09-OCT-2001; 2001US-0328044P.  
 09-OCT-2001; 2001US-0328056P.  
 12-OCT-2001; 2001US-0328849P.  
 15-OCT-2001; 2001US-0329414P.  
 17-OCT-2001; 2001US-0330142P.  
 18-OCT-2001; 2001US-0330309P.  
 22-OCT-2001; 2001US-0341058P.  
 24-OCT-2001; 2001US-0343629P.  
 29-OCT-2001; 2001US-0349575P.  
 01-NOV-2001; 2001US-0346357P.  
 17-APR-2002; 2002US-0373260P.  
 19-APR-2002; 2002US-0373815P.  
 19-APR-2002; 2002US-0373817P.  
 19-APR-2002; 2002US-0373826P.  
 19-APR-2002; 2002US-0373884P.  
 22-APR-2002; 2002US-0374977P.  
 16-MAY-2002; 2002US-0381037P.  
 16-MAY-2002; 2002US-0381038P.  
 17-MAY-2002; 2002US-0381642P.  
 28-MAY-2002; 2002US-0383656P.  
 29-MAY-2002; 2002US-0383831P.  
 25-JUN-2002; 2002US-0391335P.  
 (SMIT/) SMITHSON G.  
 (MILL/) MILLET I.  
 (PEYM/) PEYMAN J A.

PA (KEKU//) KEKUDA R.  
PA (JUJ//) JU J.  
PA (LIL//) LI L.  
PA (GUOX//) GUO X.  
PA (PAT//) PATTURAJAN M.  
PA (SPY//) SPYTEK K A.  
PA (EDIN//) EDINGER S R.  
PA (ELLE//) ELLERMAN K.  
PA (MAL//) MALYANKAR U M.  
PA (ORT//) ORT T.  
PA (GORM//) GORMAN L.  
PA (ZERH//) ZERHUSEN B D.  
PA (ANDE//) ANDERSON D W.  
PA (ZHON//) ZHONG M.  
PA (CAT//) CATTERTON E.  
PA (JIW//) JI W.  
PA (MILL//) MILLER C E.  
PA (RASI//) RASTELLI L.  
PA (STON//) STONE D J.  
PA (PENA//) PENNA C E A.  
PA (SHEN//) SHENOY S G.  
PA (SHIM//) SHIMKETS R A.  
PA (ROTH//) ROTHENBERG M E.  
PA (LEAC//) LEACH M D.  
PA (AGEE//) AGEE M L.  
PA (BERG//) BERGHS C.  
PA (DIPI//) DIPPIO V A.  
PA (EISE//) EISEN A.  
PA (GANG//) GANGOLLI E A.  
PA (RIEG//) RIEGER D K.  
PA (SPAD//) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-213931/20.  
DR N-PSDB; ADN62895.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
PS Claim 1; SEQ ID NO 92; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 250 AA;  
Query Match 100.0%; Score 40; DB 8; Length 250;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLLPLQLILL 9  
Db 2 LLLPLQLILL 10  
RESULT 8  
AAR67888  
ID AAR67888 standard; protein; 253 AA.  
XX  
AC AAR67888;  
XX  
DT 25-MAR-2003 (revised)  
DT 09-AUG-1995 (first entry)  
XX  
DE Human stratum corneum chymotrophic recombinant enzyme (SCCE).  
XX  
KW Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
KW callosities; keratosis pilaris; ichthyoses; eczema.  
XX  
OS Homo sapiens.  
XX  
PN WO9500651-A1.  
XX  
PD 05-JAN-1995.  
XX  
PF 20-JUN-1994; 94WO-IB000166.  
XX  
PR 18-JUN-1993; 93DK-00000725.  
XX  
PA (SYMB-) SYMBICOM AB.  
XX  
PI Egelrud T, Hansson L;  
XX  
DR WPI; 1995-052088/07.  
DR N-PSDB; AAQ81203.  
XX  
PT Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
PT related vectors, transformed cells and polypeptides, useful for treating  
PT skin disorders, e.g. acne or psoriasis, and for identification of  
PT specific inhibitors.  
PS Disclosure; Page 97; 137pp; English.  
XX  
CC The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
CC and skin care products, especially to treat and prevent acne, xeroderma,  
CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
CC recombinantly following mammal, insect, plant, or microorganism  
CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX  
SQ Sequence 253 AA;  
Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLLPLQLILL 9  
Db 5 LLLPLQLILL 13  
RESULT 9  
AAW05383

AAW05383 standard; protein; 253 AA.  
 AAW05383;  
 31-DEC-1996 (first entry)  
 Human amyloid precursor protein protease.  
 Amyloid precursor protein protease; Alzheimer's disease; diagnosis;  
 therapy.  
 Homo sapiens.  
 WO9631122-A1.  
 10-OCT-1996.  
 02-APR-1996; 96WO-US004294.  
 04-APR-1995; 95US-00416257.  
 (ELIL ) LILLY & CO ELI.  
 Dixon EP, Johnstone EM, Little SP;  
 WPI: 1996-464594/46.  
 N-PSDB; AAT39783.  
 New isolated human amyloid precursor protein protease - used to develop  
 prods. for the treatment or diagnosis of associated conditions, esp.  
 Alzheimer's disease.  
 Claim 1; Page 44-45; 55pp; English.  
 Human amyloid precursor protein protease (AAW05383) is involved in the  
 processing or clearance of amyloid precursor protein to form beta-amyloid  
 peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)  
 obtd. from a human lung library. Recombinant protease can be produced in  
 CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic  
 CC (partic. AV-120 host cells. It is used to develop products for the design  
 CC and testing of cpds. useful for treating or preventing conditions  
 CC associated with beta-amyloid peptide, esp. Alzheimer's disease  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLLPLQILL 9  
 Db 5 LLLPLQILL 13  
 RESULT 10  
 ABB84421  
 ID ABB84421 standard; peptide; 253 AA.  
 XX  
 AC ABB84421;  
 XX  
 DT 08-NOV-2002 (first entry)  
 XX  
 DE Human SCCE protein N-terminal fragment SEQ ID 48.  
 XX  
 KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
 KW pruritus; atopic dermatitis; eczema; acne; itch; KLUK7; enzyme.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200262135-A2.  
 XX

PD 15-AUG-2002.  
 XX  
 PF 08-FEB-2002; 2002WO-IB001300.  
 XX  
 PR 09-FEB-2001; 2001CA-02332655.  
 PR 09-FEB-2001; 2001DK-00000218.  
 XX  
 PA (EGEL/) EGELRUD T.  
 PA (HANS/) HANSSON L.  
 XX  
 PI Egelrud T, Hansson L;  
 XX  
 DR WPI: 2002-643380/69.  
 XX  
 PT Transgenic mammal or its embryo useful as model for human disease, has  
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
 PT enzyme operably linked to promoter that drives its expression in skin.  
 XX  
 PS Example 6; Page 37; 74pp; English.  
 XX  
 CC This invention describes a novel non-human transgenic mammal or mammalian  
 CC embryo having integrated within its genome, a heterologous nucleotide  
 CC sequence comprising at least a significant part of a nucleotide sequence  
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
 CC operably linked to a promoter that drives expression of heterologous SCCE  
 CC or its variant in skin. The product of the invention is useful as a model  
 CC for the study of disease with the aim of improving treatment, to relieve  
 CC or ameliorate a pathogenic condition, for development or testing of a  
 CC cosmetic or a pharmaceutical formulation, and for the development of a  
 CC diagnostic method. It can also be used as a model for a skin disease or  
 CC skin cancer. The invention is also useful for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of an  
 CC abnormal or unwanted phenotype, and for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of  
 CC inflammatory skin diseases selected from diseases consisting of epidermal  
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
 CC as a model for further studies of itch mechanisms and the testing of  
 CC potential compounds and compositions for relieve of various skin diseases  
 CC where itch is a component. This sequence represents the N-terminal  
 CC fragment of the human stratum corneum chymotryptic enzyme, SCCE  
 CC synonymous with human kallikrein 7 (KLK7), used in the development of the  
 CC transgenic mammals described in the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLLPLQILL 9  
 Db 5 LLLPLQILL 13  
 RESULT 11  
 ABB84406  
 ID ABB84406 standard; protein; 253 AA.  
 XX  
 AC ABB84406;  
 XX  
 DT 08-NOV-2002 (first entry)  
 XX  
 DE Human SCCE protein.  
 XX  
 KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
 KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200262135-A2.  
 XX 15-AUG-2002.  
 XX 08-FEB-2002; 2002WO-IB001300.  
 XX 09-FEB-2001; 2001CA-02332655.  
 PR 09-FEB-2001; 2001DK-00000218.  
 XX (EGEL/) EGELRUD T.  
 PA (HANS/) HANSSON L.  
 XX  
 PI Egelrud T, Hansson L;  
 DR WPI; 2002-643380/69.  
 DR N-PSDB; ABQ76226.  
 XX  
 PT Transgenic mammal or its embryo useful as model for human disease, has  
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
 PT enzyme operably linked to promoter that drives its expression in skin.  
 XX  
 PS Claim 10; Page 58-59; 74pp; English.  
 PS  
 CC This invention describes a novel non-human transgenic mammal or mammalian  
 CC embryo having integrated within its genome, a heterologous nucleotide  
 CC sequence comprising at least a significant part of a nucleotide sequence  
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
 CC operably linked to a promoter that drives expression of heterologous scce  
 CC or its variant in skin. The product of the invention is useful as a model  
 CC for the study of disease with the aim of improving treatment, to relieve  
 CC or ameliorate a pathogenic condition, for development or testing of a  
 CC cosmetic or a pharmaceutical formulation, and for the development of a  
 CC diagnostic method. It can also be used as a model for a skin disease or  
 CC skin cancer. The invention is also useful for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of an  
 CC abnormal or unwanted phenotype, and for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of  
 CC inflammatory skin diseases selected from diseases consisting of epidermal  
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
 CC as a model for further studies of itch mechanisms and the testing of  
 CC potential compounds and compositions for relieve of various skin diseases  
 CC where itch is a component. This sequence represents the human stratum  
 CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymous  
 CC with human kallikrein 7 (KLK7) and is used in the development of the  
 CC transgenic mammals described in the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLLPLQLL 9  
 Db 5 LLLPLQLL 13  
 RESULT 12  
 AAU82740  
 ID AAU82740 standard; protein; 253 AA.  
 XX  
 AC AAU82740;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE Amino acid sequence of novel human protease #39.  
 KW Human; protease; cancer; immune-related disorder; cardiovascular disease;  
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;

KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
 KW ocular disease; cytostatic; enzyme.  
 OS Homo sapiens.  
 XX WO200200860-A2.  
 PN 03-JAN-2002.  
 XX 26-JUN-2001; 2001WO-US020171.  
 XX 26-JUN-2000; 2000US-0214047P.  
 PR (SUGE-) SUGEN INC.  
 XX Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
 PI Charyczak G;  
 XX WPI; 2002-139913/18.  
 DR N-PSDB; ABK31782.  
 XX  
 PT Nucleic acids encoding novel human proteases, useful for useful for  
 PT treating diseases and disorders such as cancers, immune-related diseases  
 PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory  
 PT disorders.  
 PS  
 XX Claim 6; Fig 2N; 313pp; English.  
 CC The present invention relates to the isolation of novel human proteases,  
 CC and the nucleic acids encoding them. The sequences of the invention are  
 CC useful for treating diseases and disorders such as cancers (e.g. breast,  
 CC colon, lung), immune-related diseases and disorders (e.g. inflammatory  
 CC diseases and asthma), cardiovascular diseases (e.g. restenosis and  
 CC coronary thrombosis), brain or neuronal-associated diseases, metabolic  
 CC disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.  
 CC rheumatoid arthritis and psoriasis), central or peripheral nervous system  
 CC diseases, migraines, pain, sexual dysfunction, mood disorders, attention  
 CC disorders, cognitive disorders, hypotension, hypertension, psychotic  
 CC disorders, neurological disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease) and dyskinesias. The nucleic acids and polypeptides are also  
 CC useful for treating viral infections caused by human immunodeficiency  
 CC virus (HIV), and non-viral infections such as ocular disease (e.g.  
 CC glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel  
 CC human proteases of the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLLPLQLL 9  
 Db 5 LLLPLQLL 13  
 RESULT 13  
 ABU07440  
 ID ABU07440 standard; protein; 253 AA.  
 XX  
 AC ABU07440;  
 XX  
 DT 28-JAN-2003 (first entry)  
 XX  
 DE Protein differentially regulated in prostate cancer #43.  
 XX  
 KW Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX  
 OS Homo sapiens.  
 XX WO200281638-A2.  
 PN

XX PD 17-OCT-2002.

XX PF 08-APR-2002; 2002WO-US010824.

XX PR 06-APR-2001; 2001US-0281731P.

XX PR 06-APR-2001; 2001US-0281732P.

XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PI Sun Z, Jay G;

XX DR WPI; 2003-058520/05.

XX DR N-PSDB; ABX10343.

XX PT Novel genes which are differentially regulated in prostate cancer, useful for diagnosing prostate cancer in prostate tissue sample and assessing therapeutic or preventive intervention in prostate cancer patients.

XX PS Claim 1; Page 293-294; 416pp; English.

XX CC The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9

Db 5 LLLPLQILL 13

RESULT 14

ABU07471

ID ABU07471 standard; protein; 253 AA.

XX AC ABU07471;

XX

DT 28-JAN-2003 (first entry)

DE Protein differentially regulated in prostate cancer #74.

XX KW Prostate cancer; gene expression; differential regulation; molecular marker; drug target; cancer detection; cancer diagnosis; cancer staging; cancer grading; cancer assessing; cancer monitoring.

XX OS Homo sapiens.

XX PN WO200281638-A2.

XX PD 17-OCT-2002.

XX PF 08-APR-2002; 2002WO-US010824.

XX PR 06-APR-2001; 2001US-0281731P.

XX PR 06-APR-2001; 2001US-0281732P.

XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PI Sun Z, Jay G;

XX DR WPI; 2003-058520/05.

XX DR N-PSDB; ABX10375.

XX PT Novel genes which are differentially regulated in prostate cancer, useful for diagnosing prostate cancer in prostate tissue sample and assessing therapeutic or preventive intervention in prostate cancer patients.

XX PS Claim 1; Page 351; 416pp; English.

XX CC The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
| | | | |  
Db 5 LLLPLQILL 13

## RESULT 15

ABR58471  
ID ABR58471 standard; protein; 253 AA.  
XX  
AC ABR58471;  
XX  
DT 07-JUL-2003 (first entry)  
XX  
DE Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.  
XX  
KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029468-A1.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031467.  
XX  
PR 02-OCT-2001; 2001US-0327135P.  
PR 30-MAY-2002; 2002US-0384531P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Algate PA, Mannion J;  
XX  
XX WPI; 2003-372001/35.  
DR  
PT New polynucleotide and polypeptide useful for diagnosing and/or treating  
PT cancer, particularly ovarian cancer, and as a vaccine.  
XX  
PS Claim 2; Page 157-158; 169pp; English.  
XX  
CC The invention relates to a novel isolated polynucleotide. The  
CC polynucleotides of the invention have cytostatic activity, and may have a  
CC use in gene therapy, and in a vaccine. The composition and methods are  
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.  
CC The composition may also be used as a vaccine to prevent cancer. The  
CC present sequence is used in the exemplification of the invention  
XX  
SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
| | | | |  
Db 5 LLLPLQILL 13

Search completed: July 13, 2005, 17:19:51  
Job time : 77.6667 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083a-33

Perfect score: 40

Sequence: 1 LLLPLQILL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416\_seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	253	2 A53968	serine proteinase
2	36	90.0	146	2 H75201	hypothetical prote
3	34	85.0	277	2 H84314	cytochrome aa3 con
4	34	85.0	369	2 S33603	surfactant protein
5	34	85.0	371	1 JN0450	conglutinin precu
6	34	85.0	371	2 I45878	conglutinin - bovi
7	34	85.0	754	2 AE0614	probable competenc
8	34	85.0	845	2 D97163	cation transport P
9	33	82.5	196	2 G65039	hypothetical prote
10	33	82.5	282	2 E70890	hypothetical prote
11	33	82.5	370	2 AB3334	daunorubicin resis
12	33	82.5	398	2 C91063	hypothetical prote
13	33	82.5	413	2 AC0834	probable membrane
14	33	82.5	426	2 C83103	hypothetical prote
15	33	82.5	470	2 A90083	hypothetical prote
16	33	82.5	475	1 A69149	O-antigen transpor
17	33	82.5	653	2 A43632	amyloid precursor-
18	32	80.0	198	2 S25656	T-cell surface gly
19	32	80.0	235	1 RWHUT8	T-cell surface gly
20	32	80.0	238	1 LNR7MA	mannose-binding le
21	32	80.0	247	1 KYHUCM	chymase (EC 3.4.21
22	32	80.0	249	2 F91095	type III secretion
23	32	80.0	249	2 B85941	hypothetical prote
24	32	80.0	264	2 C97402	probable acyltrans
25	32	80.0	264	2 AC2620	1-acyl-sn-glycerol
26	32	80.0	266	2 JC7300	tax-responsive ele
27	32	80.0	267	2 JCA4857	hepatocarcinogenes
28	32	80.0	491	2 JCS197	stromelysin 3 (EC
29	32	80.0	492	2 A44399	stromelysin 3 (EC

RESULT 1  
A53968  
serine proteinase SCCE precursor - human  
N:Alternate names: stratum corneum chymotryptic enzyme  
C:Species: Homo sapiens (man)  
C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: A53968  
R:Hansson, L.; Stroemqvist, M.; Baekman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.  
J. Biol. Chem. 269, 19420-19426, 1994  
A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme  
A:Reference number: A53968; MUID:94308225; PMID:8034709  
A:Accession: A53968  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-253 <HAN>  
A:Cross-references: UNIPROT:P49862; GB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G532504  
C:Genetics:  
A:Gene: GDB:PRSS6; SCCE  
A:Cross-references: GDB:377730  
A:Map position: 7q35-7q35  
C:Superfamily: trypsin; trypsin homology  
F:30-245/Domain: trypsin homology <TRY>

#### ALIGNMENTS

##### RESULT 1

Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
Db 5 LLLPLQILL 13

##### RESULT 2

H75201  
hypothetical protein PAB0088 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: H75201

R:anonymous; Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc

A:Reference number: A75001

A:Accession: H75201

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-146 <KAW>

A:Cross-references: UNIPROT:Q9V2D5; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CAB4906;

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0088

C:Superfamily: Pyrococcus abyssi hypothetical protein PAB0088

Query Match 90.0%; Score 36; DB 2; Length 146;  
 Best Local Similarity 77.8%; Pred. No. 7;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
 |||||::  
 Db 67 LLLPLQIII 75

## RESULT 3

H84314  
 cytochrome aa3 controlling protein [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: H84314  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Lethauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A>Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: H84314  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-277 <STO>  
 A:Cross-references: UNIPROT:Q9HPI3; GB:AE004437; NID:g10581096; PIDN:AAG19884.1; GSPDB:C  
 C:Genetics:  
 A:Gene: ccp

Query Match 85.0%; Score 34; DB 2; Length 277;  
 Best Local Similarity 66.7%; Pred. No. 33;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
 :|||:::  
 Db 97 VLLPLQVIL 105

## RESULT 4

S33603  
 surfactant protein D - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 02-Dec-1993 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
 C:Accession: S33603  
 R:Lim, B.L.; Lu, J.; Reid, K.B.M.  
 Immunology 78, 159-165, 1993  
 A>Title: Structural similarity between bovine conglutinin and bovine lung surfactant protein D.  
 A:Reference number: S33603; MUID:93170856; PMID:8436402  
 A:Accession: S33603  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-369 <LIM>  
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology  
 F:248-367/Domain: C-type lectin homology <LCH>

Query Match 85.0%; Score 34; DB 2; Length 369;  
 Best Local Similarity 77.8%; Pred. No. 44;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
 |||||::  
 Db 2 LLLPLSVLL 10

## RESULT 5

JN0450  
 conglutinin precursor - bovine  
 N:Alternate names: C3b-binding protein  
 N:Contains: conglutinin-N  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: JN0450; JC2396; S33235; A23740; S36879; S35044; I46010; A29416; S34054

R:Suzuki, Y.; Yin, Y.P.; Makino, M.; Kurimura, T.; Wakamiya, N.  
 Biochem. Biophys. Res. Commun. 191, 335-342, 1993  
 A>Title: Cloning and sequencing of a cDNA coding for bovine conglutinin.  
 A:Reference number: JN0450; MUID:93213261; PMID:8460993  
 A:Accession: JN0450  
 A:Molecule type: mRNA  
 A:Residues: 1-371 <SUZ>  
 A:Cross-references: UNIPROT:P23805; DDBJ:D14085; NID:g285643; PIDN:BAA03170.1; PID:g28564  
 R:Kawasaki, N.; Itoh, N.; Kawasaki, T.  
 Biochem. Biophys. Res. Commun. 198, 597-604, 1994  
 A>Title: Gene organization and 5'-flanking region sequence of conglutinin: A C-type mammary protein.  
 A:Reference number: JC2396; MUID:94128104; PMID:8297370  
 A:Accession: JC2396  
 A:Molecule type: mRNA  
 A:Residues: 1-371 <K2>  
 A:Note: The authors translated the codon GAT for residues 250 and 270 as Glu  
 R:Lu, J.; Laursen, S.B.; Thiel, S.; Jensenius, J.C.; Reid, K.B.M.  
 Biochem. J. 292, 157-162, 1993  
 A>Title: The cDNA cloning of conglutinin and identification of liver as a primary site of  
 A:Reference number: S33235; MUID:93277452; PMID:7684896  
 A:Accession: S33235  
 A:Molecule type: mRNA  
 A:Residues: 1-172, 'H', 174-217, 'A', 219-271, 'V', 273-371 <LUU>  
 A:Cross-references: EMBL:X71774; NID:g395267; PIDN:CAA50665.1; PID:g395268  
 A:Experimental source: liver  
 R:Lee, Y.M.; Leiby, K.R.; Allar, J.; Paris, K.; Lerch, B.; Okarma, T.B.  
 J. Biol. Chem. 266, 2715-2723, 1991  
 A>Title: Primary structure of bovine conglutinin, a member of the C-type animal lectin fa  
 A:Reference number: A23740; MUID:91131556; PMID:1993651  
 A:Accession: A23740  
 A:Molecule type: protein  
 A:Residues: 21-209, 'S', 211-371 <LBE>  
 R:Kawasaki, N.; Yokota, Y.; Kawasaki, T.  
 Arch. Biochem. Biophys. 305, 533-540, 1993  
 A>Title: Differentiation of conglutination activity and sugar-binding activity of conglut  
 A:Reference number: S36879; MUID:93384312; PMID:8373191  
 A:Accession: S36879  
 A:Molecule type: protein  
 A:Residues: 21-54, 75-86, 'X', 88-89, 'X', 91, 'X', 93-94, 208-209, 'X', 211-227 <KAW>  
 A:Experimental source: Serum  
 R:Lu, J.; Wiedemann, H.; Holmskov, U.; Thiel, S.; Timpl, R.; Reid, K.B.M.  
 Eur. J. Biochem. 215, 793-799, 1993  
 A>Title: Structural similarity between lung surfactant protein D and conglutinin. Two di  
 A:Reference number: S35044; MUID:93358905; PMID:8354286  
 A:Accession: S35044  
 A:Molecule type: protein  
 A:Residues: 75-86, 'X', 88-89, 'X', 91, 'I' <LUA>  
 A:Experimental source: lung  
 R:Young, N.M.; Leon, M.A.  
 Biochem. Biophys. Res. Commun. 143, 645-651, 1987  
 A>Title: The carbohydrate specificity of conglutinin and its homology to proteins in the  
 A:Reference number: A29416; MUID:87184551; PMID:3566740  
 A:Contents: annotation  
 R:Malhotra, R.; Laursen, S.B.; Willis, A.C.; Sim, R.B.  
 Biochem. J. 293, 15-19, 1993  
 A>Title: Research Communication. Localization of the receptor-binding site in the collect  
 A:Reference number: S34054; MUID:93319501; PMID:8328957  
 A:Contents: annotation  
 R:Liou, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarma, T.B.; Tauber, A.I.; Sastry,  
 J. Immunol. 153, 173-180, 1994  
 A>Title: Bovine conglutinin gene exon structure reveals its evolutionary relationship to  
 A:Reference number: I46010; MUID:94267222; PMID:8207234  
 A:Accession: I46010  
 A>Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-371 <LIO>  
 A:Cross-references: EMBL:U06860; NID:g507183; PIDN:AAB60624.1; PID:g514256  
 C:Comment: This protein mediates the agglutination of erythrocytes with antibody and com  
 C:Comment: This protein is a Ca2+-dependent serum lectin specific for N-acetylglucosamin  
 C:Genetics:  
 A:Gene: CGN1  
 A:Introns: 67/1; 106/1; 142/1; 178/1; 217/1; 245/1

C;Superfamily: pulmonary surfactant protein D; C-type lectin homology  
C;Keywords: calcium binding; glycoprotein; hydroxylysine; hydroxyproline  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-371/Product: conglutinin #status predicted <MAT>  
F;46-214/Region: collagen-like  
F;75-371/Product: conglutinin-N #status predicted <MA2>  
F;248-369/Domain: C-type lectin homology <LCH>  
F;63 87,99,135,141,159,162,198,210/Binding site: carbohydrate (Lys) (covalent) #status  
F;63 87,99,135,141,159,162,198,210/Modified site: 5-hydroxylysine (Lys) #status experime  
F;78,96,108,111,129,132,147,153,171,195/Modified site: 4-hydroxyproline (Pro) #status ex  
F;337/Binding site: carbohydrate (Aan) (covalent) #status experimental

Query Match 85.0%; Score 34; DB 1; Length 371;  
Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
||||| :||  
DB 2 LLLPLSVLL 10

RESULT 6  
I45878  
conglutinin - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: I45878  
R;Liou, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarna, T.B.; Tauber, A.I.; Sastry  
Gene 141, 277-281, 1994  
A;Title: Bovine conglutinin (BC) mRNA expressed in liver: cloning and characterization o  
A;Reference number: I45878; MUID:94215917; PMID:8163202  
A;Accession: I45878  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-371 <LIO>  
A;Cross-references: UNIPROT:P23805; GB:L18871; NID:9495012; PIDN:AAA20136.1; PID:9495013  
C;Superfamily: pulmonary surfactant protein D; C-type lectin homology  
F;248-369/Domain: C-type lectin homology <LCH>

Query Match 85.0%; Score 34; DB 2; Length 371;  
Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
||||| :||  
DB 2 LLLPLSVLL 10

RESULT 7  
A80614  
probable competence-related protein STY0984 [imported] - Salmonella enterica subsp. ente  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 25-Aug-2003  
C;Accession: AE0614  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AE0614  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-754 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD05383.1; PID:gl6502146; GSPDB:GN00176  
C;Genetics:  
A;Gene: STY0984  
C;Superfamily: competence protein ComEC

Query Match 85.0%; Score 34; DB 2; Length 754;  
Best Local Similarity 77.8%; Pred. No. 90;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
||||| :||  
DB 371 LLLPLQVAL 379

RESULT 8  
D97163  
cation transport P-type ATPase CAC2137 [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: D97163  
R;Nolling, J.; Breston, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: D97163  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-845 <KUR>  
A;Cross-references: UNIPROT:Q97H76; GB:AE001437; PIDN:AAK80095.1; PID:gi5025128; GSPDB:G  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC2137

Query Match 85.0%; Score 34; DB 2; Length 845;  
Best Local Similarity 87.5%; Pred. No. 1e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLPLQILL 9  
||||| :||  
DB 670 LLLPLQILL 677

RESULT 9  
G65039  
hypothetical protein b2612 - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C;Accession: G65039  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: G65039  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-196 <BLAT>  
A;Cross-references: GB:AE000347; GB:U00096; NID:92367142; PIDN:AAC75661.1; PID:gl788965;  
A;Experimental source: strain K-12, substrain MG1655

Query Match 82.5%; Score 33; DB 2; Length 196;  
Best Local Similarity 77.8%; Pred. No. 36;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
||||| :||  
DB 123 LLLPLQILL 131

RESULT 10  
E70890  
hypothetical protein Rv1978 - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C;Accession: E70890  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A:Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: E70890  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-282 <COL>  
A:Cross-references: UNIPROT:Q53979; GB:AL022073; GB:AL123456; NID:g3256024; PIDN:CAA1785  
C:Genetics:  
A:Gene: RV1978

Query Match 82.5%; Score 33; DB 2; Length 282;  
Best Local Similarity 77.8%; Pred. No. 53;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
|||||:|  
Db 196 LLLPLHLL 204

RESULT 11  
AB3334  
daunorubicin resistance transmembrane protein [imported] - Brucella melitensis (strain 1  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AB3334  
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AB3334  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-370 <KUR>  
A:Cross-references: UNIPROT:Q8YHY9; UNIPROT:Q8FZX1; GB:AE008917; PIDN:AAL51837.1; PID:gl  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0656  
A:Map position: I

Query Match 82.5%; Score 33; DB 2; Length 370;  
Best Local Similarity 75.0%; Pred. No. 69;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 8  
|||||:|  
Db 293 ILLPLQVL 300

RESULT 12  
C91063  
hypothetical protein ECs3475 [imported] - Escherichia coli (strain O157:H7, substrain RI  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: C91063  
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: C91063  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-398 <HAY>  
A:Cross-references: UNIPROT:Q8X9C3; GB:BA000007; PIDN:BA0362946; GSPDB:G  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs3475  
C:Superfamily: hypothetical protein HI0107

Query Match 82.5%; Score 33; DB 2; Length 398;

Best Local Similarity 77.8%; Pred. No. 74;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
|||||:|  
Db 101 LLAPLQILL 109

RESULT 13  
AC0834  
probable membrane protein corB [imported] - Salmonella enterica subsp. enterica serovar 1  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AC0834  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AC0834  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-413 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD05858.1; PID:g16503833; GSPDB:GN00176  
C:Genetics:  
A:Gene: corB  
C:Superfamily: hypothetical protein HI0107

Query Match 82.5%; Score 33; DB 2; Length 413;  
Best Local Similarity 77.8%; Pred. No. 77;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
|||||:|  
Db 116 LLAPLQILL 124

RESULT 14  
C83103  
hypothetical protein PA4338 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: C83103  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:2043737; PMID:10984043  
A:Accession: C83103  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-426 <STO>  
A:Cross-references: UNIPROT:Q9HW63; GB:AE004850; NID:g9950560; PIDN:AAG0772  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4338

Query Match 82.5%; Score 33; DB 2; Length 426;  
Best Local Similarity 77.8%; Pred. No. 80;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
|||||:|  
Db 301 LLAPLQILL 309

RESULT 15  
A90083  
hypothetical protein orf470 [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: A90083  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reif  
Nature 410, 1091-1096, 2001  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; MUID:11323671; PMID:11323671  
A:Accession: A90083  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-470 <DOU>  
A:Cross-references: UNIPROT:Q9SEA5; GB:AF165818; NID:G6690603; PIDN:AAF24211.1; GSPDB:GN  
C:Genetics:  
A:Gene: orf470  
A:Map position: 1  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 82.5%; Score 33; DB 2; Length 470;  
Best Local Similarity 66.7%; Pred. No. 88;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
:|||||:  
Db 253 MLLPLEILL 261

Search completed: July 13, 2005, 17:31:28  
Job time : 15.5556 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-33

Perfect score: 40

Sequence: 1 LLLPLQILL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	66	2	O6DTY1	O6DTY1 homo sapien
2	40	100.0	253	1	KLK7_HUMAN	P49862 homo sapien
3	40	100.0	253	2	Q8N5N9	Q8N5N9 homo sapien
4	38	95.0	73	2	Q8VCA9	Q8VCA9 mus musculus
5	38	95.0	104	2	Q8R5D6	Q8R5D6 mus musculus
6	38	95.0	138	2	O6PKE2	O6PKE2 mus musculus
7	38	95.0	201	1	FK11_MOUSE	O6D1M7 mus musculus
8	38	95.0	361	1	IHA_TRIVU	O77755 trichosurus
9	37	92.5	208	2	O21527	O21527 clethrionom
10	36	90.0	146	2	O9V2D5	O9V2D5 pyrococcus
11	36	90.0	205	2	O653Z3	O653Z3 oryza sativ
12	35	87.5	181	2	O6NB96	O6NB96 rhodopseudo
13	35	87.5	1527	2	O6PPA4	O6PPA4 leishmania
14	34	85.0	162	2	Q91M24	Q91M24 lumpy skin
15	34	85.0	162	2	O91T40	O91T40 lumpy skin
16	34	85.0	277	2	Q9HP13	Q9HP13 halobacteri
17	34	85.0	369	1	PSPD_BOVIN	P35246 bos taurus
18	34	85.0	369	2	O8G3A1	O8G3A1 bos taurus
19	34	85.0	371	1	CL46_BOVIN	O8MHZ9 bos taurus
20	34	85.0	371	1	CONG_BOVIN	P23805 bos taurus
21	34	85.0	373	2	Q749Q7	Q749Q7 geobacter s
22	34	85.0	754	2	O8Z8Q2	O8Z8Q2 salmonella
23	34	85.0	754	2	O8ZQC3	O8ZQC3 salmonella
24	34	85.0	775	2	Q7N6C7	Q7N6C7 photorhabdu
25	34	85.0	845	2	Q9TH76	Q9TH76 clostridium
26	33	82.5	68	2	O6KG65	O6KG65 bacterioph
27	33	82.5	109	2	Q8BJK7	Q8BJK7 mus musculu
28	33	82.5	189	2	Q7M8B8	Q7M8B8 wolfinella s
29	33	82.5	196	2	O9X620	O9X620 salmonella
30	33	82.5	196	2	O9D226	O9D226 mus musculu
31	33	82.5	282	2	O53979	O53979 mycobacteri

32 33 82.5 282 2 O7TZ68 mycobacteri  
33 33 82.5 294 2 Q7D7M8 mycobacteri  
34 33 82.5 302 2 Q6GNY6 xenopus lae  
35 33 82.5 370 2 Q8YHY9 bruceella me  
36 33 82.5 370 2 Q8FZX1 bruceella su  
37 33 82.5 398 2 Q7ABI2 escherichia  
38 33 82.5 399 2 Q83K08 shigella fl  
39 33 82.5 413 2 O8XFY3 salmonella  
40 33 82.5 413 2 Q7CPZ5 salmonella  
41 33 82.5 420 1 YFJD\_ECOLI  
42 33 82.5 426 2 Q9HW63 pseudomonas  
43 33 82.5 435 1 YA06\_HUMAN  
44 33 82.5 470 2 O9SEK5 guillardia  
45 33 82.5 475 2 O26479 methanobact

#### ALIGNMENTS

##### RESULT 1

QSDTY1 PRELIMINARY; PRT; 66 AA.  
AC Q6DTY1  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Kallikrein 7 splice variant 3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Kishi T., Michael I.P., Diamandis E.P.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY646152; AAT66047.1; -  
SQ SEQUENCE 66 AA, 7171 MW, 82E1C392BC8222FDB CRC64;

Query Match 100.0%; Score 40; DB 2; Length 66;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9

Db 5 LLLPLQILL 13

##### RESULT 2

KLK7\_HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).  
GN Name=KLK7; Synonyms=PRSS6, SCCE;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
RC TISSUE=Skin;  
RX MEDLINE=94308225; PubMed=8034709;  
RA Hansson L., Stromqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;  
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";  
RL J. Biol. Chem. 269:19420-19426(1994).  
RN [2]  
RP SEQUENCE FROM N.A.

```

RC TISSUE=Keratinocytes;
RA Yousef G.M., Scorilas A., Diamandis E.P.;
RT "Molecular characterization, mapping and tissue expression of the
RL human stratum corneum chymotryptic enzyme gene.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN
RP SEQUENCE FROM N.A.
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA Moss P., Paepel B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
RN [4]
RN
RP SEQUENCE FROM N.A.
RA Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,
RA Wallbrandt P., Egelrud T.;
RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in
RT mice; a model for chronic itchy dermatitis.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN
RP CHARACTERIZATION
RX MEDLINE=95314630; PubMed=7794273;
RA Skytt A., Stroemqvist M., Egelrud T.;
RT "Primary substrate specificity of recombinant human stratum corneum
RL chymotryptic enzyme.";
RL Biochem. Biophys. Res. Commun. 211:586-589(1995).
CC -1- FUNCTION: May catalyze the degradation of intercellular cohesive
CC structures in the cornified layer of the skin in the continuous
CC shedding of cells from the skin surface. Specific for amino acid
CC residues with aromatic side chains in the P1 position. SCCE
CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-
CC phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the
CC activation of precursors to inflammatory cytokines.
CC -1- TISSUE SPECIFICITY: Abundantly expressed in the skin and is
CC expressed by keratinocytes in the epidermis. Very low levels are
CC also seen in the brain and kidney.
CC -1- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L33404; AAC37551.1; -
CC EMBL; AF166330; AAD49718.1; -
CC EMBL; AF243527; AAG33360.1; -
CC EMBL; AF332583; AAK69624.1; -
CC PIR; A53968; A53968.
CC HSSP; P00760; IEZX.
CC MEROPS; S01.300; -.
CC Genew; HGNC:6368; KLK7.
CC H-InvDB; HIX0015373; -.
CC MIM; 604438; -.
CC GO; GO:0008236; F:serine-type peptidase activity; TAS.
CC GO; GO:0008544; P:epidermal differentiation; TAS.
CC InterPro; IPR009003; Pept_Ser_Cys.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; Trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;
FT SIGNAL; 1 22
FT PROPEP 23 29 Activation peptide.

```

```

FT CHAIN 30 253 Kallikrein 7.
FT ACT_SITE 70 70 Charge relay system (By similarity).
FT ACT_SITE 112 112 Charge relay system (By similarity).
FT ACT_SITE 205 205 Charge relay system (By similarity).
FT DISULFID 36 137 By similarity.
FT DISULFID 55 71 By similarity.
FT DISULFID ? 239 By similarity.
FT DISULFID 144 211 By similarity.
FT DISULFID 176 190 By similarity.
FT DISULFID 201 226 By similarity.
FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

RESULT 3
Q8NSN9 PRELIMINARY; PRT; 253 AA.
AC Q8NSN9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Stratum corneum chymotryptic enzyme, preproprotein.
GN Name=KLK7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Skin;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haileh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rahas S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smailus D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN
RP SEQUENCE FROM N.A.
RX TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC EMBL; BC032005; AAH32005.1; -.
CC HSSP; P00760; IEZX.
CC GO; GO:0004263; F:chymotrypsin activity; IEA.
CC GO; GO:0008233; F:peptidase activity; IEA.
CC DR GO; GO:0004295; F:trypsin activity; IEA.
CC DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
CC DR InterPro; IPR001254; Peptidase_S1.
CC DR InterPro; IPR001314; Peptidase_S1A.

```



```

DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

RESULT 4
Q8VCA9 PRELIMINARY; PRT; 73 AA.
AC Q8VCA9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fkbp11 protein.
GN Name=Fkbp11;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the FkBP-type Pfase family.
DR EMBL; BC021345; AAH21345.1; -.
DR MGD; MGI:1913370; Fkbp11.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001179; FkBP_PPIase.
DR Pfam; PF00254; FkBP_C; 1.
DR PROSITE; PS00509; FkBP_PPIase; 1.
SQ SEQUENCE 73 AA; 7819 MW; 93E8F53399BF3C11 CRC64;

Query Match 95.0%; Score 38; DB 2; Length 73;
Best Local Similarity 88.9%; Pred. No. 7.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

RESULT 4
Q8VCA9 PRELIMINARY; PRT; 73 AA.
AC Q8VCA9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fkbp11 protein.
GN Name=Fkbp11;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021345; AAH21345.1; -.
DR MGD; MGI:1913370; Fkbp11.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001179; FkBP_PPIase.
DR Pfam; PF00254; FkBP_C; 1.
DR PROSITE; PS00509; FkBP_PPIase; 1.
SQ SEQUENCE 73 AA; 7819 MW; 93E8F53399BF3C11 CRC64;

Query Match 95.0%; Score 38; DB 2; Length 73;
Best Local Similarity 88.9%; Pred. No. 7.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

RESULT 4
Q8VCA9 PRELIMINARY; PRT; 73 AA.
AC Q8VCA9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fkbp11 protein.
GN Name=Fkbp11;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021345; AAH21345.1; -.
DR MGD; MGI:1913370; Fkbp11.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001179; FkBP_PPIase.
DR Pfam; PF00254; FkBP_C; 1.
DR PROSITE; PS00509; FkBP_PPIase; 1.
SQ SEQUENCE 73 AA; 7819 MW; 93E8F53399BF3C11 CRC64;

Query Match 95.0%; Score 38; DB 2; Length 104;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 6 LLLPLQLLL 14

RESULT 6
Q6PKB2 PRELIMINARY; PRT; 138 AA.
AC Q6PKB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fkbp11 protein.

```

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor; DOI=10.1073/pnas.242603899;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Udén T.B., Toshikiyuki S., Carninci P., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;  
 RX STRAIN=FVB/N; TISSUE=Mammary tumor;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC002311; AA02311.1; -;  
 DR GO: GO:0005665; C:extracellular space; TAS.  
 DR GO: GO:0016021; C:integral to membrane; TAS.  
 DR InterPro: IPR001179; FKBP\_PPIase.  
 DR PROSITE: PS50059; FKBP\_PPIase; 1.  
 SQ SEQUENCE 138 AA; 15105 MW; C138B8B8BFD59D CRC64;  
 Query Match 95.0%; Score 38; DB 2; Length 138;  
 Best Local Similarity 88.9%; Pred. No. 14;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LLLPLQLLL 9  
 DB 6 LLLPLQLLL 14  
 RESULT 7  
 ID FK11 MOUSE STANDARD; PRT; 201 AA.  
 AC Q9DLM7; Q9CRE4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FK506 binding protein l1 precursor (FC 5.2.1.8) (Peptidyl-prolyl cis-  
 DE trans isomerase) (PPIase) (Rotamase) (19 kDa FK506-binding protein)  
 DE (FKBP-19).  
 GN Name=FKBP11;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nitaide I., Oato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Mateuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gasterland T., Gariboldi M., Giasi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltata K., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Hiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Havaishizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Udén T.B., Toshikiyuki S., Carninci P., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: PPIases accelerate the folding of proteins during  
 CC protein synthesis.  
 CC -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline  
 CC (omega=0).  
 CC -!- SIMILARITY: Belongs to the FKBP-type PPIase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AK003331; BAB22719.1; -;  
 CC EMBL: AK019132; BAB31559.1; -;  
 CC EMBL: BC037596; AAH37596.1; -;  
 CC HSSP: P20071; ITCO.  
 CC MGD: MGI:1913370; Fkbp11.  
 DR InterPro: IPR001179; FKBP\_PPIase.  
 DR Pfam: PF00254; FKBP\_C; 1.  
 DR PROSITE: PS50059; FKBP\_PPIase; 1.  
 KW isomerase; Rotamase; Signal.  
 FT SIGNAL 1 27 Potential.  
 FT CHAIN 28 201 FK506 binding protein 11.

```

FT DOMAIN 57 144 PPIase, PKBP-type.
FT CONFLICT 53 53 S -> P (in Ref. 1; BAB31559).
FT CONFLICT 198 198 S -> R (in Ref. 1; BAB31559).
SQ SEQUENCE 201 AA; 22137 MW; 94D955C57264BD82 CRC64;

Query Match 95.0%; Score 38; DB 1; Length 201;
Best Local Similarity 88.9%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9
Db 6 LLLPLQLLL 14

RESULT 8
ID IHA_TRIUV STANDARD; PRT; 361 AA.
AC O77755;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Inhibin alpha chain precursor.
GN Name=INHA;
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99027340; PubMed=9801457;
RA Vamontfort D., Fidler A.E., Heath D.A., Lawrence S.B., Tisdall D.J.,
RA Greenwood P.J., McNatty K.;
RT "cDNA sequence analysis, gene expression and protein localisation of
RT the inhibin alpha subunit of Australian brushtail possum (Trichosurus
RT vulpecula).";
RL J. Mol. Endocrinol. 21:141-152(1998).
CC -!- FUNCTION: Inhibins and activins inhibit and activate,
CC respectively, the secretion of follitropin by the pituitary gland.
CC Inhibins/activins are involved in regulating a number of diverse
CC functions such as hypothalamic and pituitary hormone secretion,
CC gonadal hormone secretion, germ cell development and maturation,
CC erythroid differentiation, insulin secretion, nerve cell survival,
CC embryonic axial development or bone growth, depending on their
CC subunit composition. Inhibins appear to oppose the functions of
CC activins.
CC -!- SUBUNIT: Dimeric, linked by one or more disulfide bonds. Inhibin A
CC is a dimer of alpha and beta-A. Inhibin B is a dimer of alpha and
CC beta-B.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF033340; AAC63945.1; -.
DR GO; GO:0005576; C:extracellular; ISS.
DR GO; GO:0017106; F:activin inhibitor activity; ISS.
DR GO; GO:0005125; F:cytokine activity; ISS.
DR GO; GO:0008083; F:growth factor activity; ISS.
DR GO; GO:0005515; F:hormone activity; ISS.
DR GO; GO:0005515; F:protein binding; ISS.
DR GO; GO:0007050; P:cell cycle arrest; ISS.
DR GO; GO:0030154; P:cell differentiation; ISS.
DR GO; GO:0008151; P:cell growth and/or maintenance; ISS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. .; ISS.
DR GO; GO:0007267; P:cell-cell signaling; ISS.
DR GO; GO:0030218; P:erythrocyte differentiation; ISS.
DR GO; GO:0042541; P:hemoglobin biosynthesis; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.

DR GO; GO:0045578; P:negative regulation of B-cell differentiation; ISS.
DR GO; GO:0045786; P:negative regulation of cell cycle; ISS.
DR GO; GO:0046882; P:negative regulation of follicle-stimulating. .; ISS.
DR GO; GO:0045077; P:negative regulation of interferon-gamma bio. .; ISS.
DR GO; GO:0045650; P:negative regulation of macrophage different. .; ISS.
DR GO; GO:0042326; P:negative regulation of macrophage different. .; ISS.
DR GO; GO:0007399; P:negative regulation of phosphorylation; ISS.
DR GO; GO:0001541; P:neurogenesis; ISS.
DR GO; GO:0001541; P:ovarian follicle development; ISS.
DR GO; GO:0046881; P:positive regulation of follicle-stimulating. .; ISS.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR Pfam; PF00019; TGF_beta; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein; Growth factor; Hormone; Signal.
FT SIGNAL 1 21 Potential.
FT PROPEP 22 230 By similarity.
FT CHAIN 231 361 Inhibin alpha chain.
FT DISULFID 260 323 By similarity.
FT DISULFID 289 358 By similarity.
FT DISULFID 293 360 By similarity.
FT DISULFID 322 322 Interchain (By similarity).
FT CARBOHYD 48 48 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 144 144 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 266 266 N-linked (GlcNAc. .) (Potential).
SQ SEQUENCE 361 AA; 38945 MW; D661CDF93CDAA87D CRC64;

Query Match 95.0%; Score 38; DB 1; Length 361;
Best Local Similarity 88.9%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9
Db 4 LLLPLQLLL 12

RESULT 9
ID O21527 PRELIMINARY; PRT; 208 AA.
AC O21527;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
GN Name=ND4;
OS Clethrionomys gapperi (Southern red-backed vole).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
OC Clethrionomys.
OX NCBI_TaxID=56223;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98152303; PubMed=9491603;
RA Engel S.R., Hogan K.M., Taylor J.F., Davis S.K.;
RT "Molecular systematics and paleobiogeography of the South American
RT sigmodontine rodents.";
RL Mol. Biol. Evol. 15:35-49(1998).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; U83808; AAB87168.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. .; IEA.
DR InterPro; IPR003918; NADH_oxred4.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR PRINTS; PR01437; NUOXDRDTASE4.

```

```
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 208 AA; 208 208 8AF1788697AED6A2 CRC64;

Query Match 92.5%; Score 37; DB 2; Length 208;
Best Local Similarity 88.9%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
   ||:|||||
Db 98 LLVPLQILL 106

RESULT 10
Q9V2D5 PRELIMINARY; PRT; 146 AA.
AC Q9V2D5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FAB0088;
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Poch O., Priet J., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Welssenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi."
RL Mol. Microbiol. 47:1495-1512(2003).
DR EMBL; AJ248283; CAB49063.1; -.
DR PIR; H75201;
DR InterPro; IPR008537; DUF819.
DR Pfam; PF03684; DUF819; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 146 AA; 16092 MW; 7182941371258C1F CRC64;

Query Match 90.0%; Score 36; DB 2; Length 146;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
   ||:|||||
Db 67 LLLPLQILL 75

RESULT 11
Q653Z3 PRELIMINARY; PRT; 205 AA.
AC Q653Z3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Endonuclease V protein-like.
GN Name=OSJNB0065C04.45;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OSJNB0065C04."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004744; BAD45874.1; -.

KW Endonuclease.
SQ SEQUENCE 205 AA; 22733 MW; 6FBFE7178F6C9EFF CRC64;

Query Match 90.0%; Score 36; DB 2; Length 205;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
   ||:|||||
Db 176 LLLPLQILL 184

RESULT 12
Q6NB96 PRELIMINARY; PRT; 181 AA.
AC Q6NB96;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RPA0932;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y. Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris."
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572595; CAE26376.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 181 AA; 19632 MW; CDB44D8210E565D1 CRC64;

Query Match 87.5%; Score 35; DB 2; Length 181;
Best Local Similarity 87.5%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLPLQILL 9
   ||:|||||
Db 24 LLPLQVLL 31

RESULT 13
Q6PPA4 PRELIMINARY; PRT; 1527 AA.
AC Q6PPA4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE P-glycoprotein D.
GN Name=PGPD;
OS Leishmania tarentolae (Saurleishmania tarentolae).
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania;
OC Lizard Leishmania.
OX NCBI_TaxID=5689;
RN [1]
RP SEQUENCE FROM N.A.
RA Leprohon P.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY589043; AAT02643.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity; IEA. coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
```

DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC\_membrane.1.  
 DR InterPro; IPR001140; ABC\_TM\_transpct.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00664; ABC\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00929; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1527 AA; 167943 MW; E943681FFC675AAE CRC64;

Query Match 87.5%; Score 35; DB 2; Length 1527;  
 Best Local Similarity 77.8%; Pred. No. 5.7e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
 | | | | |  
 DB 424 LFLPLQVLL 432

## RESULT 14

Q91M24 PRELIMINARY; PRT; 162 AA.  
 ID Q91M24;  
 AC Q91M24;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE LSDV010 LAP/PHD-finger protein.  
 GN Name=LSDV010; Synonyms=LD010.  
 OS Lumpy skin disease virus (LSDV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Capripoxvirus.  
 OX NCBI\_TaxID=59509;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21329495; PubMed=11435593;  
 RX DOI=10.1128/JVI.75.15.7122-7130.2001;  
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak A., Kutish G.F., Rock D.L.;  
 RT "Genome of lumpy skin disease virus.";  
 RL J. Virol. 75:7122-7130(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=22710395; PubMed=12827464;  
 RA Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Abolnik C., Lu Z.,  
 RA Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;  
 RT "Comparative sequence analysis of the South African vaccine strain and  
 two virulent field isolates of Lumpy skin disease virus.";  
 RL Arch. Virol. 148:1335-1356(2003).  
 DR EMBL; AF325528; AAK84971.1;  
 DR EMBL; AF409137; AAN02576.1;  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00744; RINGV; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 SQ SEQUENCE 162 AA; 18835 MW; 0EA24745C3818222 CRC64;

Query Match 85.0%; Score 34; DB 2; Length 162;  
 Best Local Similarity 88.9%; Pred. No. 1e+02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
 | | | | |  
 DB 90 LLLPLTILL 98

## RESULT 15

Q91T40 PRELIMINARY; PRT; 162 AA.  
 ID Q91T40;  
 AC Q91T40;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Hypothetical protein (LAP/PHD-finger protein).  
 GN Name=LW010;  
 OS Lumpy skin disease virus (LSDV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Capripoxvirus.  
 OX NCBI\_TaxID=59509;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Neethling;  
 RA Stipinovich C., Vreede F.T., Kara P.D., Wallace D.B., Nel L.H.,  
 RA Viljoen G.J.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22710395; PubMed=12827464;  
 RA Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Abolnik C., Lu Z.,  
 RA Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;  
 RT "Comparative sequence analysis of the South African vaccine strain and  
 two virulent field isolates of Lumpy skin disease virus.";  
 RL Arch. Virol. 148:1335-1356(2003).  
 DR EMBL; AF336128; AAK43550.1;  
 DR EMBL; AF409138; AAN02734.1;  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00744; RINGV; 1.  
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.  
 SQ SEQUENCE 162 AA; 18782 MW; 5F914A4080F729EE CRC64;

Query Match 85.0%; Score 34; DB 2; Length 162;  
 Best Local Similarity 88.9%; Pred. No. 1e+02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
 | | | | |  
 DB 90 LLLPLTILL 98

Search completed: July 13, 2005, 17:29:20  
 Job time : 64.2222 secs

**This page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-33

Perfect score: 40

Sequence: 1 LLLPLQILL 9

Scoring table: BLOSUM62

:: Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	3	US-09-502-600-33
2	40	100.0	9	4	US-09-918-243-33
3	40	100.0	253	2	US-08-557-146-2
4	40	100.0	253	2	US-08-824-874-3
5	40	100.0	253	2	US-09-154-344-2
6	40	100.0	253	3	US-08-930-188-2
7	40	100.0	253	3	US-09-210-084-3
8	40	100.0	253	4	US-09-764-762-3
9	40	100.0	253	5	PCT-US96-04294-2
10	40	100.0	265	4	US-09-949-016-7716
11	38	95.0	812	4	US-09-489-039A-12075
12	36	90.0	9	3	US-09-502-600-35
13	36	90.0	9	3	US-09-502-600-36
14	36	90.0	9	4	US-09-918-243-35
15	36	90.0	9	4	US-09-918-243-36
16	33	82.5	190	1	US-08-339-152A-19
17	33	82.5	190	2	US-08-007-999B-6
18	33	82.5	190	2	US-08-689-276A-6
19	33	82.5	447	4	US-09-252-991A-24312
20	33	82.5	634	1	US-08-339-152A-17
21	33	82.5	653	1	US-08-339-152A-16
22	33	82.5	653	2	US-08-007-999B-3
23	33	82.5	653	2	US-08-689-276A-3
24	32	80.0	21	3	US-08-753-007A-24
25	32	80.0	21	3	US-09-398-496-24
26	32	80.0	22	3	US-08-977-378-22
27	32	80.0	235	1	US-07-940-605A-12

28 32 80.0 235 2 US-08-690-096-12 Sequence 12, Appl  
29 32 80.0 295 4 US-09-270-767-39889 Sequence 39889, A  
30 32 80.0 295 4 US-09-270-767-55106 Sequence 55106, A  
31 32 80.0 302 4 US-09-393-634-37 Sequence 37, Appl  
32 32 80.0 376 3 US-08-751-512-8 Sequence 8, Appl  
33 32 80.0 436 4 US-09-489-039A-11785 Sequence 11785, A  
34 32 80.0 492 1 US-07-794-393-4 Sequence 4, Appl  
35 32 80.0 492 1 US-08-001-711-4 Sequence 4, Appl  
36 32 80.0 837 4 US-09-489-039A-12646 Sequence 12646, A  
37 32 80.0 1306 3 US-08-989-299-7 Sequence 7, Appl  
38 32 80.0 1306 4 US-09-407-427-7 Sequence 7, Appl  
39 31 77.5 129 4 US-09-513-999C-4177 Sequence 4177, Ap  
40 31 77.5 134 4 US-09-621-976-3902 Sequence 3902, Ap  
41 31 77.5 142 4 US-09-248-796A-21399 Sequence 21399, A  
42 31 77.5 180 4 US-09-270-767-42872 Sequence 42872, A  
43 31 77.5 226 4 US-09-869-388-10 Sequence 10, Appl  
44 31 77.5 227 4 US-09-869-388-6 Sequence 6, Appl  
45 31 77.5 270 2 US-08-773-368-1 Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-09-502-600-33  
; Sequence 33, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 33  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 5-13 of the SCCE protein  
US-09-502-600-33

Query Match 100.0%; Score 40; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
Db 1 LLLPLQILL 9

RESULT 2  
US-09-918-243-33  
; Sequence 33, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 33  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens





APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-154-344-2

Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.6; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
DB 5 LLLPLQILL 13

RESULT 6  
US-08-930-188-2  
Sequence 2, Application US/08930188  
Patent No. 6093397  
GENERAL INFORMATION:  
APPLICANT: Dixon, Eric P.  
ADDRESSEE: Johnstone, Edward M.  
APPLICANT: Little, Shella P.  
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
RELATED NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930,188  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/416,257  
FILING DATE: 04-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Blalock, Donna K.  
REGISTRATION NUMBER: 38,082  
REFERENCE/DOCKET NUMBER: X9239  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-277-1090  
TELEFAX: 317-276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-930-188-2

Query Match 100.0%; Score 40; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.6; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLPLQILL 9  
DB 5 LLLPLQILL 13  
RESULT 7  
US-09-210-084-3  
Sequence 3, Application US/09210084  
Patent No. 6197511  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,084  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,874  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0252 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 532504  
US-09-210-084-3

Query Match 100.0%; Score 40; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.6; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9  
DB 5 LLLPLQILL 13

RESULT 8  
US-09-764-762-3  
Sequence 3, Application US/09764762  
Patent No. 6472195  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5

;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;/ STREET: 3174 Porter Drive  
;/ CITY: Palo Alto  
;/ STATE: CA  
;/ COUNTRY: USA  
;/ ZIP: 94304  
;/  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Diskette  
;/ COMPUTER: IBM Compatible  
;/ OPERATING SYSTEM: DOS  
;/ SOFTWARE: FastSeq for Windows Version 2.0  
;/  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/09/764,762  
;/ FILING DATE: 16-Jan-2001  
;/ CLASSIFICATION: <Unknown>  
;/  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: 09/210,084  
;/ FILING DATE: <Unknown>  
;/  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Billings, Lucy J.  
;/ REGISTRATION NUMBER: 36,749  
;/ REFERENCE/DOCKET NUMBER: PF-0252 US  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 415-855-0555  
;/ TELEFAX: 415-845-4166  
;/  
;/ INFORMATION FOR SEQ ID NO: 3:  
;/  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 253 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ IMMEDIATE SOURCE:  
;/ LIBRARY: GenBank  
;/ CLONE: 532504  
;/  
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
;/

US-09-764-762-3

Query Match 100.0%; Score 40; DB 4; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9  
Db 5 LLLPLQLLL 13

RESULT 9  
PCT-US96-04294-2  
; Sequence 2, Application PC/TUS9604294  
; GENERAL INFORMATION:  
; APPLICANT: Dixon, Eric P.  
; APPLICANT: Johnstone, Edward M.  
; APPLICANT: Little, Sheila P.  
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
; TITLE OF INVENTION: RELATED NUCLEIC ACIDS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/04294  
; FILING DATE:  
; CLASSIFICATION:

;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: US 08/416,257  
;/ FILING DATE: 04-APR-1995  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Blalock, Donna K.  
;/ REGISTRATION NUMBER: 38,082  
;/ REFERENCE/DOCKET NUMBER: X9239  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 317-277-1090  
;/ TELEFAX: 317-276-3861  
;/ INFORMATION FOR SEQ ID NO: 2:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 253 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ PCT-US96-04294-2

Query Match 100.0%; Score 40; DB 5; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9  
Db 5 LLLPLQLLL 13

RESULT 10  
US-09-949-016-7716  
; Sequence 7716, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR APPLICATION NUMBER: 2000-04-14  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7716  
; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-7716

Query Match 100.0%; Score 40; DB 4; Length 265;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9  
Db 17 LLLPLQLLL 25

RESULT 11  
US-09-489-039A-12075  
; Sequence 12075, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27

```

: PRIOR APPLICATION NUMBER:  US 60/117,747
:
: PRIOR FILING DATE:  1999-01-29
:
: NUMBER OF SEQ ID NOS:  14342
:
: SEQ ID NO 12075
:
: LENGTH: 812
:
: TYPE: PRT
:
: ORGANISM:  Klebsiella pneumoniae
:
: US-09-4849-0139A-12075

```

Query Match 95.0%; Score 38; DB 4; Length 812;  
Best Local Similarity 88.9%; Pred. No. 30;  
Matches 8; Conservative 1; Mismatches 0; Indels

Qy 1 LLLPLQILL 9  
429 LLLPLQILL 437

```

RESULT 12
US-09-502-600-35
; Sequence 35, Application US/09502600A
; Patent NO. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; - OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-502-600-35

```

Query Match 90.0%; Score 36; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred.No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels

Qy 2 LLPLQILL 9  
| | | | |  
Db 1 LLPLQILL 8

```

RESULT 13
US-09-502-600-36
; Sequence 36, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 4-12 of the SCCE protein
; US-09-502-600-36

```

Query Match 90.0%; Score 36; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;

```

Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1      LLLPLQL 8
          |||||
Db      2      LLLPLOIL 9

```

```

RESULT 14
US-09-918-243-35
; Sequence 35, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-918-243-35

```

Query Match	90.0%	Score 36;	DB 4;	Length 9;
Best Local Similarity	100.0%	Pred. No. 4.1e+05;		
Matches 8;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

QY 2 LLPLQILL 9  
| | | | |  
Db 1 LLPLQILL 8

```

RESULT 15
US-09-918-243-36
; Sequence 36, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 4-12 of the SCCE protein
US-09-918-243-36

```

Query Match	90.0%	Score 36;	DB 4;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 4.1e+05;		
Matches ..	8;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;

Qy 1 LLLPLQIL 8  
|||  
2 LLLPLQIL 9  
db

Search completed: July 13, 2005, 17:34:26  
Job time : 20.4444 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds  
(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-33  
Perfect score: 40  
Sequence: 1 LLLPLQLLL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	9 US-09-918-243-33	Sequence 33, Appl
2	40	100.0	9	9 US-09-905-083-33	Sequence 33, Appl
3	40	100.0	9	15 US-10-372-521-33	Sequence 33, Appl
4	40	100.0	9	15 US-10-831-075-33	Sequence 33, Appl
5	40	100.0	198	15 US-10-262-511-96	Sequence 96, Appl
6	40	100.0	250	15 US-10-262-511-96	Sequence 92, Appl
7	40	100.0	253	9 US-09-888-615-98	Sequence 98, Appl
8	40	100.0	253	9 US-09-764-762-3	Sequence 3, Appl
9	40	100.0	253	14 US-10-071-214-2	Sequence 2, Appl
10	40	100.0	253	14 US-10-071-214-48	Sequence 48, Appl
11	40	100.0	253	14 US-10-264-283-90	Sequence 90, Appl

12	40	100.0	253	15 US-10-295-027-498	Sequence 498, Appl
13	40	100.0	253	15 US-10-173-999-48	Sequence 48, Appl
14	40	100.0	253	16 US-10-408-765A-639	Sequence 639, Appl
15	40	100.0	253	16 US-10-643-795A-95	Sequence 95, Appl
16	40	100.0	253	17 US-10-948-518-95	Sequence 95, Appl
17	40	100.0	253	17 US-10-868-490A-1	Sequence 1, Appl
18	40	100.0	257	15 US-10-344-394-38	Sequence 38, Appl
19	38	95.0	70	15 US-10-424-599-207950	Sequence 207950
20	36	90.0	9	9 US-09-918-243-35	Sequence 35, Appl
21	36	90.0	9	9 US-09-918-243-36	Sequence 36, Appl
22	36	90.0	9	9 US-09-905-083-35	Sequence 35, Appl
23	36	90.0	9	9 US-09-905-083-36	Sequence 36, Appl
24	36	90.0	9	15 US-10-372-521-35	Sequence 35, Appl
25	36	90.0	9	15 US-10-372-521-36	Sequence 35, Appl
26	36	90.0	9	16 US-10-831-075-35	Sequence 35, Appl
27	36	90.0	9	16 US-10-831-075-36	Sequence 36, Appl
28	36	90.0	187	16 US-10-437-963-110905	Sequence 110905
29	35	87.5	99	16 US-10-425-115-348070	Sequence 348070
30	35	87.5	173	16 US-10-437-963-119305	Sequence 119305
31	34	85.0	23	9 US-09-864-761-46097	Sequence 46097, A
32	34	85.0	61	16 US-10-425-115-218429	Sequence 218429
33	34	85.0	187	15 US-10-425-114-57128	Sequence 57128, A
34	34	85.0	201	10 US-09-956-622A-39	Sequence 39, Appl
35	34	85.0	229	16 US-10-425-115-227327	Sequence 227327
36	34	85.0	369	18 US-10-820-155-60	Sequence 60, Appl
37	34	85.0	369	18 US-10-820-155-64	Sequence 64, Appl
38	34	85.0	369	18 US-10-820-155-74	Sequence 74, Appl
39	34	85.0	371	18 US-10-820-155-3	Sequence 3, Appl
40	34	85.0	371	18 US-10-820-155-4	Sequence 24, Appl
41	34	85.0	371	18 US-10-820-155-24	Sequence 54, Appl
42	34	85.0	371	18 US-10-820-155-54	Sequence 54, Appl
43	34	85.0	845	15 US-10-282-122A-51875	Sequence 51875, A
44	33	82.5	46	15 US-10-424-599-265948	Sequence 265948
45	33	82.5	58	16 US-10-437-963-156728	Sequence 156728

## ALIGNMENTS

RESULT 1  
US-09-918-243-33  
; Sequence 33, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 33  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN.  
; OTHER INFORMATION: Residues 5-13 of the SCCE protein  
US-09-918-243-33

Query Match	100.0%	Score 40;	DB 9;	Length 9;
Best Local Similarity	100.0%	Pred. No. 1.6e+06;		
Matches	9;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
Qy	1 LLLPLQLLL 9			
Db	1 LLLPLQLLL 9			
RESULT 2				

```
US-09-905-083-33
; Sequence 33, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-905-083-33

Query Match      100.0%; Score 40; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLLPLQILL 9
Db      1 LLLPLQILL 9

RESULT 3
US-10-372-521-33
; Sequence 33, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-10-372-521-33

Query Match      100.0%; Score 40; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLLPLQILL 9
Db      1 LLLPLQILL 9

RESULT 4
US-10-831-075-33
; Sequence 33, Application US/10831075
; Publication No. US20040224891A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
```

```
; FILE REFERENCE: D6223CIP/C/D/CIP3
; CURRENT APPLICATION NUMBER: US/10/831,075
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 10/372,521
; PRIOR FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-10-831-075-33

Query Match      100.0%; Score 40; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLLPLQILL 9
Db      1 LLLPLQILL 9

RESULT 5
US-10-262-511-96
; Sequence 96, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
```

```
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 96
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-96

Query Match      100.0%; Score 40; DB 15; Length 198;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 6
US-10-262-511-92
; Sequence 92, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerkhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: 2003-05-28
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
```

```
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 92
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-92

Query Match      100.0%; Score 40; DB 15; Length 250;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 2 LLLPLQILL 10

RESULT 7
US-09-888-615-98
; Sequence 98, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAN, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Query Match      100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 8
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKEIN
; NUMBER OF SEQUENCES: 5
```

;/  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;/ STREET: 3174 Porter Drive  
;/ CITY: Palo Alto  
;/ STATE: CA  
;/ COUNTRY: USA  
;/ ZIP: 94304  
;/  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Diskette  
;/ COMPUTER: IBM Compatible  
;/ OPERATING SYSTEM: DOS  
;/ SOFTWARE: FastSeq for Windows Version 2.0  
;/  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/09/764,762  
;/ FILING DATE: 16-Jan-2001  
;/ CLASSIFICATION: <unknown>  
;/  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: 09/210,084  
;/ FILING DATE: <unknown>  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Billings, Lucy J.  
;/ REGISTRATION NUMBER: 36,749  
;/ REFERENCE/DOCKET NUMBER: PF-0252 US  
;/  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 415-855-0555  
;/ TELEFAX: 415-845-4166  
;/  
;/ INFORMATION FOR SEQ ID NO: 3:  
;/  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 253 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/  
;/ IMMEDIATE SOURCE:  
;/ LIBRARY: GenBank  
;/ CLONE: 532504  
;/  
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-764-762-3

Query Match 100.0%; Score 40; DB 9; Length 253;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
| | | | |  
Db 5 LLLPLQILL 13

RESULT 9  
US-10-071-214-2  
; Sequence 2, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGERLUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-071-214-2

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
| | | | |  
Db 5 LLLPLQILL 13

RESULT 10  
US-10-071-214-48  
; Sequence 48, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGERLUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 48  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Deduced amino acid sequence from the C-terminal part of SCCE from  
US-10-071-214-48

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
| | | | |  
Db 5 LLLPLQILL 13

RESULT 11  
US-10-264-283-90  
; Sequence 90, Application US/10264283  
; Publication No. US20030144494A1  
; GENERAL INFORMATION:  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.590  
; CURRENT APPLICATION NUMBER: US/10/264,283  
; CURRENT FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 90  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-283-90

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9  
| | | | |  
Db 5 LLLPLQILL 13

RESULT 12  
US-10-295-027-498  
; Sequence 498, Application US/10295027



```
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 13
US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US2004000563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
```

```
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-48

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 14
US-10-408-765A-639
; Sequence 639, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 639
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-639

Query Match      100.0%; Score 40; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 15
US-10-643-795A-95
; Sequence 95, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026F1-US
; CURRENT APPLICATION NUMBER: US/10/643,795A
```

```
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 95
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-643-795A-95
```

```
Query Match      100.0%; Score 40; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLLPLQILL 9
        |||||
Db      5 LLLPLQILL 13
```

Search completed: July 13, 2005, 18:23:37  
Job time : 68.8889 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083A-34

Perfect score: 47

Sequence: 1 VLWNERWTL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	47	100.0	9	4	AAE08239	Human str
2	47	100.0	9	8	ADR68795	Human str
3	47	100.0	97	6	ADA05740	Human NOV
4	47	100.0	97	8	ADN62904	Human NOV
5	47	100.0	136	4	ABG23378	Novel hum
6	47	100.0	181	6	ADA05738	Human NOV
7	47	100.0	181	8	ADN62902	Human NOV
8	47	100.0	198	6	ADA05736	Human NOV
9	47	100.0	198	8	ADN62900	Human NOV
10	47	100.0	224	6	ADA05744	Human NOV
11	47	100.0	224	8	ADN62908	Human NOV
12	47	100.0	225	4	AAE98502	Human Str
13	47	100.0	247	6	ADA05742	Human NOV
14	47	100.0	247	8	ADN62906	Human NOV
15	47	100.0	250	6	ADA05732	Human NOV
16	47	100.0	250	8	ADN62896	Human NOV
17	47	100.0	252	6	ADA05734	Human NOV
18	47	100.0	252	8	ADN62898	Human NOV
19	47	100.0	253	2	AAE67888	Human str
20	47	100.0	253	2	AAW05383	Human any
21	47	100.0	253	5	ABB84421	Human SCC
22	47	100.0	253	5	ABB84406	Human SCC
23	47	100.0	253	5	AAU82740	Amino aci
24	47	100.0	253	6	ABU07440	Protein d
25	47	100.0	253	6	ABU07471	Protein d

26	47	100.0	253	6	ABR58471	Human str
27	47	100.0	253	7	ADB80484	Ovarian c
28	47	100.0	253	7	ADJ68833	Human hea
29	47	100.0	253	7	ADN39180	Cancer/an
30	47	100.0	253	8	ADL06515	Human tum
31	47	100.0	253	8	ADN04182	Antipsori
32	47	100.0	253	8	ADR72880	Human ova
33	47	100.0	257	3	AAE21326	Human HSC
34	43	91.5	9	4	AAE08245	Human str
35	43	91.5	9	4	AAE08255	Human str
36	43	91.5	9	8	ADR68801	Human str
37	43	91.5	9	8	ADR68811	Human str
38	40	85.1	243	5	ABB84419	Bovine SC
39	40	85.1	249	5	ABB84420	Porcine S
40	39	83.0	212	2	AAE28590	Human Pac
41	39	83.0	228	7	AAE39994	Human adi
42	39	83.0	228	8	ADN04134	Antipsori
43	39	83.0	228	8	ADP03097	Human ine
44	39	83.0	228	8	ADP23853	PRO polyp
45	39	83.0	250	2	AAE05421	Human adi

## ALIGNMENTS

## RESULT 1

AAE08239  
ID AAE08239 standard; peptide; 9 AA.

XX AC AAE08239;

XX XX 01-NOV-2001 (first entry)

XX DE Human stratum corneum chymotrypsin enzyme peptide #3 (residues 58-66).

XX KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;  
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;  
KW antisense therapy; malignant hyperplasia.

XX OS Homo sapiens.

XX PN WO200159158-A1.

XX PD 16-AUG-2001.

XX PF 07-FEB-2001; 2001WO-US003977.

XX PR 11-FEB-2000; 2000US-00502600.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'brien TJ;

XX DR WPI; 2001-514676/56.

XX PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX PS Claim 25; Page 103; 127pp; English.

XX CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

SQ Sequence 9 AA;

Query Match 100.0%; Score 47; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
| | | | | | | |  
Db 1 VLVNERWVL 9

RESULT 2

AD68795  
ID AD68795 standard; peptide; 9 AA.

XX AC AD68795;  
XX

DT 02-DEC-2004 (first entry)

DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:34.  
XX  
KW serine protease; stratum corneum chymotrytic enzyme; SCCE;  
KW immune response; ovarian cancer; lung cancer; prostate cancer;  
KW pancreatic cancer; colon cancer.  
XX

OS Homo sapiens.

XX PN WO2004075723-A2.  
XX

XX PD 10-SEP-2004.

XX PF 20-FEB-2004; 2004WO-US005134.

XX PR 21-FEB-2003; 2003US-00372521.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'brien TJ, Cannon MJ, Santin A;

XX DR WPI; 2004-653294/63.

XX PT Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for  
PT vaccinating an individual against SCCE, and in monitoring and developing  
PT immunotherapies for ovarian and other malignancies.

XX PS Claim 5; SEQ ID NO 34; 117pp; English.

CC The invention relates to a novel method for vaccinating an individual  
CC against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating  
CC an individual with a SCCE peptide, which elicits an immune response in  
CC the individual. A peptide of the invention acts as a stratum corneum  
CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
CC an individual against SCCE, particularly an individual having, suspected  
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
CC an individual, such as ovarian, breast, lung, colon, prostate, or  
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
CC peptides are also useful in the monitoring and development of  
CC immunotherapies for ovarian and other malignancies. The present sequence  
CC represents a peptide fragment of serine protease SCCE (stratum corneum  
CC chymotrytic enzyme).

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 47; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
| | | | | | | |  
Db 1 VLVNERWVL 9

RESULT 3

ADA05740  
ID ADA05740 standard; protein; 97 AA.

XX ADA05740;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV18e protein SEQ ID NO:100.  
XX  
KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029424-A2.  
XX  
XX 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US01373.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339286P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 22-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 28-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.  
XX (CURA-) CURAGEN CORP.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
XX Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
XX Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton B;  
XX Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
XX Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
XX Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2003-381626/36.  
XX N-P8DB; ADA05739.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
XX preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
XX cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
XX pharmacogenomics.  
XX  
XX Claim 1; Page 171; 586pp; English.  
XX  
XX The present invention describes NOVX proteins, where X can be 1 to 55

(e.g. NOV1). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid molecule which encodes a NOVX protein of the invention; (4) a vector comprising the nucleic acid molecule described above; (5) a cell comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity or of latency or predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide. NOVX sequences have antidiabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian and antipalemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a syndrome associated with a human disease. The polypeptide or the nucleic acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence represents a human NOVX from the present invention.

Sequence 97 AA;

Query Match 100.0%; Score 47; DB 6; Length 97;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
| | | | | | | |  
Db 39 VLVNERWVL 47

RESULT 4

ID ADN62904 standard; protein; 97 AA.

AC ADN62904;

DT 01-JUL-2004 (first entry)

DE Human NOV18e.

KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease; anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorder; dyslipidaemia; metabolic syndrome X; wasting disorder.

OS Homo sapiens.

PN US2004038223-A1.

PD 26-FEB-2004.

PF 01-OCT-2002; 2002US-00262511.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
XX (SMIT//) SMITHSON G.  
PA (MILL//) MILLET I.  
PA (PEYM//) PEYMAN J. A.  
PA (KEKU//) KEKUDA R.  
PA (JUJU//) JU J.  
PA (LILL//) LI L.  
PA (GUOX//) GUO X.  
PA (PATT//) PATTURAJAN M.  
PA (SPYT//) SPYTEK K. A.  
PA (EDIN//) EDINGER S. R.  
PA (ELLE//) ELLERMAN K.  
PA (MALY//) MALYANKAR U. M.  
PA (ORTT//) ORT T.  
PA (GORM//) GORMAN L.  
PA (ZERH//) ZERHUSEN B. D.  
PA (ANDE//) ANDERSON D. W.  
PA (ZHON//) ZHONG M.  
PA (CATT//) CATTERTON E.  
PA (JIWW//) JI W.  
PA (MILL//) MILLER C. E.  
PA (RST//) RASTELLI L.  
PA (STON//) STONE D. J.  
PA (PENA//) PENNA C. E. A.  
PA (SHEN//) SHENOY S. G.  
PA (SHIM//) SHIMKETS R. A.  
PA (ROTH//) ROTHENBERG M. E.  
PA (LEAC//) LEACH M. D.  
PA (AGEE//) AGEE M. L.  
PA (BERG//) BERGHS C.  
PA (DIP//) DIPIPPO V. A.  
PA (EISE//) EISEN A.  
PA (GANG//) GANGOLLI E. A.  
PA (RIEG//) RIEGER D. K.  
PA (SPAD//) SPADERNA S. K.  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spyttek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2004-213931/20.  
DR N-PSDB; ADN62903.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing, diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 100; 395pp; English.  
 PS  
 CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX  
 SQ Sequence 97 AA;

Query Match 100.0%; Score 47; DB 8; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
 |||||  
 DB 39 VLVNERWVL 47

RESULT 5  
 ABG23378  
 ID ABG23378 standard; protein; 136 AA.

XX  
 AC ABG23378;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #23369.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS87565.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

Claim 20; SEQ ID NO 53737; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 136 AA;

Query Match 100.0%; Score 47; DB 4; Length 136;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
 |||||  
 DB 64 VLVNERWVL 72

RESULT 6

ADA05738

ID ADA05738 standard; protein; 181 AA.

AC ADA05738;

XX 06-NOV-2003 (first entry)

DE Human NOV18d protein SEQ ID NO:98.

KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

PN WO2003029424-A2.

PD 10-APR-2003.

PF 02-OCT-2002; 2002WO-US031373.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

12-OCT-2001; 2001US-0328849P.  
 15-OCT-2001; 2001US-0329414P.  
 17-OCT-2001; 2001US-0330142P.  
 18-OCT-2001; 2001US-0330309P.  
 22-OCT-2001; 2001US-03341058P.  
 24-OCT-2001; 2001US-0339266P.  
 24-OCT-2001; 2001US-0343629P.  
 29-OCT-2001; 2001US-0349575P.  
 01-NOV-2001; 2001US-0346357P.  
 17-APR-2002; 2002US-037260P.  
 19-APR-2002; 2002US-0373815P.  
 19-APR-2002; 2002US-0373817P.  
 19-APR-2002; 2002US-0373826P.  
 22-APR-2002; 2002US-0373884P.  
 22-APR-2002; 2002US-0374977P.  
 16-MAY-2002; 2002US-0381037P.  
 16-MAY-2002; 2002US-0381038P.  
 17-MAY-2002; 2002US-0381042P.  
 17-MAY-2002; 2002US-0381642P.  
 28-MAY-2002; 2002US-0383656P.  
 29-MAY-2002; 2002US-0383831P.  
 25-JUN-2002; 2002US-0391335P.  
 01-OCT-2002; 2002US-00262511.  
 (CURA-) CURAGEN CORP.  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 Ott T, Gorman L, Zerkus BN, Anderson BW, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CBA, Shenoy SG;  
 Shimkete RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiPippo VA;  
 Eisen AJ, Gangolli EA, Kieger DK, Spaderna SK;  
 WPI; 2003-381626/36.  
 N-PSDB; ADA05737.  
 New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 pharmacogenomics.  
 Claim 1; Page 171; 586pp; English.  
 The present invention describes NOVX proteins, where X can be 1 to 55  
 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 described above and a carrier; (2) a kit comprising, in one or more  
 containers, the composition described above; (3) an isolated nucleic acid  
 molecule which encodes a NOVX protein of the invention; (4) a vector  
 comprising the nucleic acid molecule described above; (5) a cell  
 comprising the above vector; (6) an antibody that immunospecifically  
 binds to the polypeptide described above; (7) methods for determining the  
 presence or amount of the above polypeptide or nucleic acid molecule in a  
 sample; (8) methods for determining the presence of or predisposition to  
 a disease associated with altered levels of expression of the above  
 polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 method of identifying an agent that binds to the polypeptide described  
 above; (10) a method for identifying a potential therapeutic agent for  
 use in treating a pathology that is related to an aberrant expression or  
 aberrant physiological interactions of the polypeptide; (11) a method of  
 screening for a modulator of activity or of latency or predisposition to  
 a pathology associated with the polypeptide; (12) a method for modulating  
 the activity of the polypeptide described above; (13) methods of treating  
 or preventing a pathology associated with the above polypeptide in a  
 mammal; and (14) a method for producing the above polypeptide. NOVX  
 sequences have antidiabetic, anorectic, antibacterial, virucide,  
 immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 and antiplatelet activities, and can be used in gene therapy. The  
 polypeptide is useful in manufacturing a medicament for treating a  
 syndrome associated with a human disease. The polypeptide or the nucleic  
 acid molecule may be used to diagnose, treat or prevent metabolic  
 disorders such as diabetes or obesity, infections, cachexia, cancer,  
 neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.  
 XX  
 SQ Sequence 181 AA;  
 Query Match 100.0%; Score 47; DB 6; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 2.3; 0; Mismatches 0; Gaps 0;  
 Matches 9; Conservative 0;  
 Qy 1 VLVNERWVL 9  
 |||||  
 Db 39 VLVNERWVL 47  
 RESULT 7  
 ADN62902  
 ID ADN62902 standard; protein; 181 AA.  
 XX  
 AC ADN62902;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human NOV18d.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 29-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-037260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX  
 (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.

PA (KEKU//) KEKUDA R.  
 PA (JUJ//) JU J.  
 PA (LILL//) LI L.  
 PA (GUOX//) GUO X.  
 PA (PATT//) PATTURAJAN M.  
 PA (SPYT//) SPYTEK K A.  
 PA (EDIN//) EDINGER S R.  
 PA (ELLE//) ELLERMAN K.  
 PA (MALY//) MALYANKAR U M.  
 PA (ORTT//) ORT T.  
 PA (GORM//) GORMAN L.  
 PA (ZERH//) ZERHUSEN B D.  
 PA (ANDE//) ANDERSON D W.  
 PA (ZHON//) ZHONG M.  
 PA (CATT//) CATTERTON E.  
 PA (JIW//) JI W.  
 PA (MILL//) MILLER C E.  
 PA (RAST//) RASTELLI L.  
 PA (STON//) STONE D J.  
 PA (PENA//) PENNA C E A.  
 PA (SHEN//) SHENOY S G.  
 PA (SHIM//) SHIMKETS R A.  
 PA (ROTH//) ROTHENBERG M E.  
 PA (LEAC//) LEACH M D.  
 PA (AGEE//) AGEE M L.  
 PA (BERG//) BERGHS C.  
 PA (DIP//) DIPIPPO V A.  
 PA (EISE//) EISEN A.  
 PA (GANG//) GANGOLLI E A.  
 PA (RIEG//) RIEGER D K.  
 PA (SPAD//) SPADERNA S K.  
 XX  
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CE, Shenoy SG;  
 PI Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 DR WPI: 2004-213931/20.  
 DR N-PSDB: ADN62901.  
 XX  
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 XX  
 PS Claim 1; SEQ ID NO 98; 395pp; English.  
 XX  
 CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disease, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX  
 SQ Sequence 181 AA;  
 Query Match 100.0%; Score 47; DB 8; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 VLVNERWVL 9  
 Db 39 VLVNERWVL 47  
 |||||  
 RESULT 8  
 ADA05736  
 ID ADA05736 standard; protein; 198 AA.  
 XX  
 AC ADA05736;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV18c protein SEQ ID NO:96.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX  
 OS Homo sapiens.

XX  
 PN WO2003029424-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 02-OCT-2002; 2002WO-US031373.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 28-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI



PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DM, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2003-381626/36.  
DR N-PSDB; ADA05735.  
DR N-PSDB; ADA05735.  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX Claim 1; Page 170; 586pp; English.  
XX The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOVX). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, norectic, antihypertensive, antiparkinsonian  
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
CC and antilipidemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX SQ Sequence 198 AA;  
Query Match 100.0%; Score 47; DB 6; Length 198;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLVNERWVL 9  
DB 58 VLVNERWVL 66  
|||||||  
RESULT 9  
ADN62900  
ID ADN62900 standard; protein; 198 AA.  
XX AC ADN62900;  
XX 01-JUL-2004 (first entry)  
XX Human NOV18c.  
DE human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder; X;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.  
XX Homo sapiens.  
OS US2004038223-A1.  
PN 26-FEB-2004.  
XX 01-OCT-2002; 2002US-00262511.  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
XX (SMIT//) SMITHSON G.  
PA (MILL//) MILLET I.  
PA (PEYM//) PEYMAN J A.  
PA (KEKU//) KESKUDA R.  
PA (JUNJ//) JU J.  
PA (LILL//) LI L.  
PA (GUOX//) GUO X.  
PA (PATT//) PATTURAJAN M.  
PA (SPYT//) SPYTEK K A.  
PA (EDIN//) EDINGER S R.  
PA (ELLE//) ELLERMAN K.  
PA (MALY//) MALYANKAR U M.  
PA (ORTT//) ORT T.  
PA (GORM//) GORMAN L.  
PA (ZERH//) ZERHUSEN B D.  
PA (ANDE//) ANDERSON D W.  
PA (ZHON//) ZHONG M.  
PA (CATT//) CATTERTON E.  
PA (JIWW//) JI W.  
PA (MILL//) MILLER C E.  
PA (RAST//) RASTELLI L.  
PA (STON//) STONE D J.  
PA (PENNA//) PENNA C E A.  
PA (SHEN//) SHENOY S G.  
PA (SHIM//) SHIMKETS R A.  
PA (ROTH//) ROTHENBERG M E.  
PA (LEAC//) LEACH M D.  
PA (AGEE//) AGEE M L.  
PA (BERG//) BERGHs C.  
PA (DIP//) DIPIPPO V A.

PA (EISEN) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-213931/20.  
DR N-PSDB; ADN62899.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
XX Claim 1; SEQ ID NO 96; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 198 AA;  
Query Match 100.0%; Score 47; DB 8; Length 198;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 VLVNERWYL 9  
| | | | | | | | | |  
Db 58 VLVNERWYL 66  
RESULT 10  
ADA05744  
ID ADA05744 standard; protein; 224 AA.  
XX  
AC ADA05744;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV18g protein SEQ ID NO:104.  
XX  
XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipaemic; gene therapy; human disease;  
KW

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031373.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
XX WPI; 2003-381626/36.  
DR N-PSDB; ADA05743.  
XX  
PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Claim 1; Page 172; 586pp; English.  
XX  
CC The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, neurotropic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipidemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 224 AA;

Query Match 100.0%; Score 47; DB 6; Length 224;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9

Db 44 VLVNERWVL 52

RESULT 11

ADN62908

ID ADN62908 standard; protein; 224 AA.

XX AC ADN62908;

XX DT 01-JUL-2004 (first entry)

XX DE Human NOV189.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

XX OS Homo sapiens.

XX XX US2004038223-A1.

XX PD 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 09-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENNA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIPV/) DIPPIO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2004-213931/20.

DR N-PSDB; ADN62907.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 104; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules

CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 224 AA;

Query Match 100.0%; Score 47; DB 8; Length 224;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
DB 44 VLVNERWVL 52  
|||||

RESULT 12

AAB98502  
ID AAB98502 standard; protein; 225 AA.

AC AAB98502;

DT 03-AUG-2001 (first entry)

DE Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.

KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;

KW tumour antigen-derived gene 15; serine protease;

KW Stratum Corneum Chymotryptic Enzyme; SCCE.

OS Homo sapiens.

PN WO200129056-A1.

PD 26-APR-2001.

PF 20-OCT-2000; 2000WO-US029095.

PR 20-OCT-1999; 99US-00421213.

PA (UYAR-) UNIV ARKANSAS.

PI O'Brien TV, Tanimoto H;

DR WPI; 2001-381031/40.

PT Novel extracellular serine protease, termed tumor antigen-derived gene 15  
PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,  
PT treatment, prevention of cancer, particularly breast, ovarian cancer.

PS Example 10; Fig 1; 130pp; English.

XX The present invention relates to human tumour antigen-derived gene 15

CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500). TADG-  
CC 15 is an extracellular serine protease. It was found that TADG-15 is over

CC -expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20

CC residues that lack TADG-15 protease activity are useful for vaccinating  
CC an individual against TADG-15, having, suspected of having or at risk of  
CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic  
CC or therapeutic target in cancer. The present sequence was used in a  
CC sequence homology alignment with the catalytic domain of TADG-15

XX Sequence 225 AA;

Query Match 100.0%; Score 47; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
DB 30 VLVNERWVL 38  
|||||

RESULT 13

ADA05742

ID ADA05742 standard; protein; 247 AA.

XX ADA05742;

DT 06-NOV-2003 (first entry)

DE Human NOV18f protein SEQ ID NO:102.

KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytostatic; nootropic; neuroprotective;

KW antiparkinsonian; antilipaeamic; gene therapy; human disease;

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

XX immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

PN WO2003029424-A2.

PD 10-APR-2003.

PF 02-OCT-2002; 2002WO-US031373.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX

(CURA-) CURAGEN CORP.  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Patturajan M, Spyttek KA, Edinger SR, Ellerman K, Malvankar UM;  
 Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 WPI; 2003-381626/36.  
 N-PSDB; ADA05741.  
 New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 pharmacogenomics.  
 Claim 1; Page 172; 586pp; English.  
 The present invention describes NOVX proteins, where X can be 1 to 55  
 (e.g. NOVX). Also described: (1) a composition comprising a polypeptide  
 described above and a carrier; (2) a kit comprising, in one or more  
 containers, the composition described above; (3) an isolated nucleic acid  
 molecule which encodes a NOVX protein of the invention; (4) a vector  
 comprising the nucleic acid molecule described above; (5) a cell  
 comprising the above vector; (6) an antibody that immunospecifically  
 binds to the polypeptide described above; (7) methods for determining the  
 presence or amount of the above polypeptide or nucleic acid molecule in a  
 sample; (8) methods for determining the presence of or predisposition to  
 a disease associated with altered levels of expression of the above  
 polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 method of identifying an agent that binds to the polypeptide described  
 above; (10) a method for identifying a potential therapeutic agent for  
 use in treating a pathology that is related to an aberrant expression or  
 aberrant physiological interactions of the polypeptide; (11) a method of  
 screening for a modulator of activity or of latency or predisposition to  
 a pathology associated with the polypeptide; (12) a method for modulating  
 the activity of the polypeptide described above; (13) methods of treating  
 or preventing a pathology associated with the above polypeptide in a  
 mammal; and (14) a method for producing the above polypeptide. NOVX  
 sequences have anti-diabetic, anorectic, antibacterial, virucide,  
 immunomodulator, cytoskeletal, neurotropic, neuroprotective, antiparkinsonian  
 and anti-lipase activities, and can be used in gene therapy. The  
 polypeptide is useful in manufacturing a medicament for treating a  
 syndrome associated with a human disease. The polypeptide or the nucleic  
 acid molecule may be used to diagnose, treat or prevent metabolic  
 disorders such as diabetes or obesity, infections, cachexia, cancer,  
 neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 disease, immune disorders, haematopoietic disorders and various  
 dyslipidaemias. The nucleic acids can also be used as hybridisation  
 probes, in chromosome mapping, tissue typing, preventive medicine and  
 pharmacogenomics. The present sequence represents a human NOVX from the  
 present invention.  
 Sequence 247 AA;  
 Query Match 100.0%; Score 47; DB 6; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VLVNERWYL 9  
 Db 44 VLVNERWYL 52  
 RESULT 14  
 ADN62906  
 ID ADN62906 standard; protein; 247 AA.  
 XX  
 AC ADN62906;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX

DE Human NOV18f.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 12-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX  
 (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALVANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENNA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHINKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.

PA (AGEE/) AGEE M L.  
PA (BERG/) BERGHS C.  
PA (DIP/) DIPIPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-2113931/20.  
DR N-PSDB; ADN62905.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
XX Claim 1; SEQ ID NO 102; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 247 AA;  
  
Query Match 100.0%; Score 47; DB 8; Length 247;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VLVNRRWVL 9  
Db 44 VLVNRRWVL 52  
  
RESULT 15  
ADA05732  
ID ADA05732 standard; protein; 250 AA.  
XX  
AC ADA05732;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV18a protein SEQ ID NO:92.  
XX

KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031373.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2003-381626/36.  
DR N-PSDB; ADA05731.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX Claim 1; Page 169-170; 586pp; English.  
PS  
XX The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
 CC and antilipidemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX  
 SQ Sequence 250 AA;

Query Match 100.0%; Score 47; DB 6; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 3.3; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
 |||||  
 Db 55 VLVNERWVL 63

Search completed: July 13, 2005, 17:19:52  
 Job time : 77.6667 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083A-34

Perfect score: 47

Sequence: 1 VLVNERWVL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	253	2	A53968
2	39	83.0	246	1	DBHU
3	38	80.9	232	1	KQPG
4	38	80.9	260	2	I56559
5	38	80.9	417	2	T25594
6	37	78.7	250	2	T01779
7	37	78.7	250	2	S31384
8	37	78.7	258	1	S44184
9	37	78.7	259	1	S49129
10	37	78.7	1421	2	T05892
11	36	76.6	72	2	P82825
12	36	76.6	258	2	A45161
13	36	76.6	426	2	D42519
14	36	76.6	426	2	T37408
15	36	76.6	426	2	T21666
16	36	76.6	426	2	T28563
17	36	76.6	426	2	C36850
18	36	76.6	516	2	T47075
19	35	74.5	178	2	T25570
20	35	74.5	231	2	S31778
21	35	74.5	242	2	S31776
22	35	74.5	242	2	S31775
23	35	74.5	242	2	S43489
24	35	74.5	247	1	A25852
25	35	74.5	257	1	JC2479
26	35	74.5	258	2	G02959
27	35	74.5	258	2	I36945
28	35	74.5	258	2	I36947
29	35	74.5	274	2	I47078

30 35 74.5 274 2 E70472  
31 35 74.5 282 2 A83757  
32 35 74.5 329 1 HPDG  
33 35 74.5 345 2 I36941  
34 35 74.5 346 2 I36942  
35 35 74.5 347 1 HPHU1  
36 35 74.5 347 2 G00006  
37 35 74.5 348 1 HPHUR  
38 35 74.5 349 2 I36944  
39 35 74.5 406 1 HPHU2  
40 35 74.5 410 2 T46034  
41 35 74.5 416 1 KFO  
42 35 74.5 421 1 S11674  
43 35 74.5 453 2 G96695  
44 35 74.5 505 2 C86443  
45 35 74.5 539 2 S54045

## ALIGNMENTS

### RESULT 1

A53968

serine proteinase SCCE precursor - human

N:Alternate names: stratum corneum chymotryptic enzyme

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A53968

R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A:Reference number: A53968; MUID:94308225; PMID:8034709

A:Accession: A53968

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <HAN>

A:Cross-references: UNIPROT:P49862; GB:L33404; NID:9521214; PIDN:AAC37551.1; PID:9532504

C:Genetics: GDB:PRSS6; SCCE

A:Cross-references: GDB:377730

A:Map position: 7q35-7q35

C:Superfamily: trypsin; trypsin homology

F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 47; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9

Db 58 VLVNERWVL 66

### RESULT 2

DBHU

complement factor D (EC 3.4.21.46) precursor [validated] - human (fragment)

N:Alternate names: adipsin; C3 convertase activator

C:Species: Homo sapiens (man)

C>Date: 28-Aug-1985 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004

C:Accession: A40197; A00936; A60571; S66645

R:White, R.T.; Damm, D.; Hancock, N.; Rosen, B.S.; Lowell, B.B.; Usher, P.; Flier, J.S.;

J. Biol. Chem. 267, 9210-9213, 1992

A:Title: Human adipsin is identical to complement factor D and is expressed at high level

A:Reference number: A40197; MUID:92250520; PMID:1374388

A:Accession: A40197

A:Molecule type: mRNA

A:Residues: 1-246 <WHI>

A:Cross-references: UNIPROT:P00746; GB:M84526

R:Niemann, M.A.; Bhow, A.S.; Bennett, J.C.; Volanakis, J.B.

Biochemistry 23, 2482-2486, 1984

A:Title: Amino acid sequence of human D of the alternative complement pathway.

A:Reference number: A00936; MUID:85000441; PMID:6383466

A:Accession: A00936

A:Molecule type: protein  
A:Residues: 19-44,'G',46-51,'Q',53-75,'TH',78,'P',80-83,'XXXITIE',90-172,86-91,185-235,'  
A>Note: a few residues were assigned from the previously published sequence of Reid et al  
R,Miyata, T.; Oda, O.; Inagi, R.; Sugiyama, S.; Miyama, A.; Maeda, K.; Nakashima, I.; Ya  
Mol. Immunol. 27, 637-644, 1990  
A:Title: Molecular and functional identification and purification of complement component  
A:Reference number: A60571; MUID:90370044; PMID:2395435  
A:Accession: A60571  
A:Molecule type: protein  
A:Residues: 19-20,'XX',23-27,'XX',30-31,'XX',34,'X',36-40 <MIY>  
R,Baake, N.; Holtkamp, U.; Hoerl, W.H.; Tschesche, H.  
FEBS Lett. 371, 300-302, 1995  
A:Title: Inhibition of degranulation of human polymorphonuclear leukocytes by complement  
A:Reference number: S66645; MUID:96013156; PMID:7556615  
A:Accession: S66645  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 19-44,'C',46-48 <BAL>  
C:Comment: Factor D cleaves factor B when the latter is complexed with factor C3b, activ  
C:Genetics:  
A:Gene: GDB:DF  
A:Cross-references: GDB:132645; OMIM:134350  
A:Map position: Xpter-Xqter  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase  
F:1-18/Domain: complement sequence #status predicted <SIG>  
F:19-246/Product: complement factor D (fragment) #status experimental <MAT>  
F:19-241/Domain: trypsin homology <TRY>  
F:19-60/141-207,172-188,197-222/Disulfide bonds: #status predicted  
F:44-140,141-207,172-188,197-222/Disulfide bonds: #status predicted  
F:59-160,161-201/Active site: His, Asp, Ser #status predicted

Query Match 83.0%; Score 39; DB 1; Length 246;  
Best Local Similarity 88.9%; Pred. No. 7.8;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
||| |||||  
DB 47 VLVNERWVL 55

RESULT 3  
KOPG  
tissue kallikrein (EC 3.4.21.35), pancreatic - pig (tentative sequence)  
N/Alternate names: glandular kallikrein; kininogenin  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 24-Apr-1984 #sequence\_revision 31-Dec-1993 #text\_change 31-Mar-2000  
C:Accession: A00938; A92895  
R:Tschesche, H.; Mair, G.; Godec, G.; Fiedler, F.; Ehret, W.; Hirschauer, C.; Lemon, M.;  
Adv. Exp. Med. Biol. 120, 245-260, 1979  
A:Title: The primary structure of porcine glandular kallikreins.  
A:Reference number: A90015  
A:Accession: A00938  
A:Molecule type: protein  
A:Residues: 1-49,'GWL',53-134,'D',136-156,'H',158,'B',160-224,'B',226-232 <TSC>  
A>Note: the residue identified as 225-Asx is bound to carbohydrate; therefore, we have s  
R:Node, W.; Chen, Z.; Bartels, K.; Kutzbach, C.; Schmidt-Kastner, G.; Bartunik, H.  
J. Mol. Biol. 164, 237-282, 1983  
A:Title: Refined 2 angstrom X-ray crystal structure of porcine pancreatic kallikrein A,  
t, structure and its comparison with bovine trypsin.  
A:Reference number: A92895; MUID:83189107; PMID:6551452  
A:Contents: X-ray crystallography, 2 angstroms  
A:Accession: A92895  
A:Molecule type: protein  
A:Residues: 1-224,'B',226-232 <BOD>  
C:Comment: The protein consists of two chains, A and B, held together by disulfide bonds  
C:Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release 1  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: glycoprotein; hydrolase; pancreas; serine proteinase  
F:1-224/Domain: trypsin homology <TRY>  
F:1-80/Product: tissue kallikrein chain A #status experimental <MPTA>  
F:79-82/Region: autolysis loop  
F:81-232/Product: tissue kallikrein chain B #status experimental <MPFB>  
F:7-144,26-42,121-190,155-169,180-205/Disulfide bonds: #status experimental

F:41,89,184/Active site: His, Asp, Ser #status experimental  
F:78/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:225/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental

Query Match 80.9%; Score 38; DB 1; Length 232;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
||| |||||  
DB 29 VLVNPKWVL 37

RESULT 4  
neuropsin - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I56559  
R:Chen, Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Nishino  
J. Neurosci. 15, 5088-5097, 1995  
A:Title: Expression and activity-dependent changes of a novel limbic-serine protease gene  
A:Reference number: I56559; MUID:95348817; PMID:7623137  
A:Accession: I56559  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-260 <RES>  
A:Cross-references: UNIPROT:Q61955; GB:D30785; NID:g1648847; PIDN:BAA06451.1; PID:g102005  
C:Superfamily: trypsin; trypsin homology  
F:33-252/Domain: trypsin homology <TRY>

Query Match 80.9%; Score 38; DB 2; Length 260;  
Best Local Similarity 77.8%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
||| |||||  
DB 61 VLVGDRWVL 69

RESULT 5  
T25594  
hypothetical protein C32E8.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25594  
R:Gatung, S.  
submitted to the EMBL Data Library, February 1997  
A:Description: The sequence of C. elegans cosmid C32E8.  
A:Reference number: Z20056  
A:Accession: T25594  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-417 <GAT>  
A:Cross-references: UNIPROT:P91123; EMBL:U88308; PIDN:AAB42324.1; GSPDB:GN00019; CESP:C3;  
A:Experimental source: strain Bristol N2; clone C32E8  
C:Genetics:  
A:Gene: CESP:C32E8.6  
A:Map position: 1  
A:Introns: 22/3; 51/2; 92/1; 297/1; 354/1

Query Match 80.9%; Score 38; DB 2; Length 417;  
Best Local Similarity 85.7%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERW 7  
||| |||||  
DB 219 VLVNERW 225

RESULT 6  
T01779  
trypsin (EC 3.4.21.4) - plaice

C:Species: Pleuronectes platessa (plaice)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 16-Jul-1999  
C:Accession: T01779

R:Leaver, M.J.; George, S.G.  
submitted to the EMBL Data Library, March 1996

A:Reference number: Z14422

A:Accession: T01779

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-250 <LEA>

A:Cross-references: EMBL:X56744; NID:g1213630; PID:G64240

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:23-242/Domain: trypsin homology <TRY>

Query Match 78.7%; Score 37; DB 2; Length 250;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9

DB 50 VLINNQWVL 58

#### RESULT 7

S41384

trypsin (EC 3.4.21.4) - plaice

C:Species: Pleuronectes platessa (plaice)

C:Date: 02-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004

C:Accession: S31384

R:Leaver, M.J.; George, S.G.

submitted to the EMBL Data Library, November 1990

A:Reference number: S30567

A:Accession: S31384

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-250 <LEA>

A:Cross-references: UNIPROT:P35034; EMBL:X56744

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:23-242/Domain: trypsin homology <TRY>

Query Match 78.7%; Score 37; DB 2; Length 250;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9

DB 50 VLINNQWVL 58

#### RESULT 8

S41184

chymotrypsin (EC 3.4.21.1) 2 precursor - African malaria mosquito

N:Alternate names: chymotrypsin-like proteinase ANCHYM2

C:Species: Anopheles gambiae (African malaria mosquito)

C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C:Accession: S44184; S49130

R:Mueller, H.M.; Catteruccia, F.; Crisanti, A.

submitted to the EMBL Data Library, April 1994

A:Reference number: S44184

A:Accession: S44184

A:Molecule type: DNA

A:Residues: 1-258 <MUE>

A:Cross-references: UNIPROT:Q17025; EMBL:Z32645; NID:g474026; PID:CAA83567.1; PID:g474028

R:Mueller, H.

submitted to the EMBL Data Library, November 1992

A:Description: Cloning of blood meal induced serine protease genes of the mosquito Anopheles gambiae

A:Reference number: S35412

A:Accession: S49130

A:Molecule type: DNA

A:Residues: 1-165, 'R', 167-193, 'FPD', 197-258 <MUE>

A:Cross-references: EMBL:Z18888; NID:g509416; PIDN:CAA79326.1; PID:g509417

C:Genetics:

A:Introns: 78/2; 209/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-32/Domain: activation peptide #status predicted <PRO>

F:33-258/Product: chymotrypsin 2 #status predicted <MAT>

F:33-250/Domain: trypsin homology <TRY>

F:59-75,182-198,208-232/Bisulfide bonds: #status experimental

F:74,119,212/Active site: His, Asp, Ser #status predicted

Query Match 78.7%; Score 37; DB 1; Length 258;  
Best Local Similarity 75.0%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVNERWVL 9

DB 63 LLNDRWVL 70

#### RESULT 9

S49129

chymotrypsin (EC 3.4.21.1) 1 precursor - African malaria mosquito

N:Alternate names: chymotrypsin-like proteinase ANCHYM1

C:Species: Anopheles gambiae (African malaria mosquito)

C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C:Accession: S49129; S44185

R:Mueller, H.

submitted to the EMBL Data Library, November 1992

A:Description: Cloning of blood meal induced serine protease genes of the mosquito Anopheles gambiae

A:Reference number: S35412

A:Accession: S49129

A:Molecule type: mRNA

A:Residues: 1-259 <MUE>

A:Cross-references: UNIPROT:Q27289; EMBL:Z18887; NID:g509414; PIDN:CAA79325.1; PID:g50941

A:Experimental source: strain Suakoko

R:Mueller, H.M.; Catteruccia, F.; Crisanti, A.

submitted to the EMBL Data Library, April 1994

A:Description: An Anopheles gambiae locus containing the sequences of two closely related

A:Reference number: S44184

A:Accession: S44185

A:Molecule type: DNA

A:Residues: 1-259 <MUE>

A:Cross-references: EMBL:Z32645; NID:g474026; PIDN:CAA83568.1; PID:g474028

A:Experimental source: strain Suakoko

C:Genetics:

A:Introns: 78/2; 209/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-32/Domain: activation peptide #status predicted <PRO>

F:33-259/Product: chymotrypsin 1 #status predicted <MAT>

F:33-250/Domain: trypsin homology <TRY>

F:59-75,182-198,208-232/Bisulfide bonds: #status experimental

F:74,119,212/Active site: His, Asp, Ser #status predicted

Query Match 78.7%; Score 37; DB 1; Length 259;  
Best Local Similarity 75.0%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVNERWVL 9

DB 63 LLNDRWVL 70

#### RESULT 10

T05892

hypothetical protein F6H11.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 09-Apr-1999 #sequence\_revision 09-Apr-1999 #text\_change 13-Aug-1999

C:Accession: T05892

R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Bancroft, I.; Mewer

submitted to the Protein Sequence Database, April 1998  
 A;Reference number: Z15456  
 A;Accession: T05892  
 A;Molecule type: DNA

A;Residues: 1-1421 <BEV>  
 A;Cross-references: EMBL:AL021684  
 A;Experimental source: cultivar Columbia; BAC clone F6H11  
 C;Genetics:

A;Map position: 5  
 A;Intron: 72/3; 142/3; 227/2; 344/3; 372/3; 941/2; 1016/3; 1038/3; 1055/1; 1095/2; 1159  
 A;Note: F6H1.110

Query Match 78.7%; Score 37; DB 2; Length 1421;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
 ||:||||:  
 Db 1208 VLANKRWVI 1216

## RESULT 11

F82825  
 hypothetical protein XF0279 [imported] - Xylella fastidiosa (strain 9a5c)  
 C;Species: Xylella fastidiosa  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A;Reference number: A82515; PMID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: F82825

A;Status: preliminary  
 A;Molecule type: DNA

A;Cross-references: UNIPROT:Q9PGL9; GB:AE003881; GB:AE003849; NID:g9105093; PIDN:AAF8309  
 A;Experimental source: strain 9a5c  
 R;Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 Briones, M.R.S.; Bueno, M.R.P.; Canargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.F.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
 A;Authors: Martins, B.M.F.; Matsukuma, A.V.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A;Reference number: A59328

A;Contents: annotation  
 C;Genetics:

A;Gene: XF0279

Query Match 76.6%; Score 36; DB 2; Length 72;  
 Best Local Similarity 62.5%; Pred. No. 7.7;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWV 8  
 ||:||||:  
 Db 50 ILMNERWL 57

## RESULT 12

A45161  
 serine proteinase (EC 3.4.21.-) RNK-Met-1 - rat (fragment)  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 09-Jul-2004  
 C;Accession: A45161  
 R;Smyth, M.J.; Wilcrot, T.; Trapani, J.A.; Ottaway, K.S.; Sowder, R.; Henderson, L.E.;  
 J. Biol. Chem. 267, 24418-24425, 1992  
 A;Title: Purification and cloning of a novel serine protease, RNK-Met-1, from the granu

A;Reference number: A45161; PMID:93077530; PMID:1447189  
 A;Accession: A45161  
 A;Status: preliminary  
 A;Molecule type: mRNA

A;Residues: 1-258 <SMY>  
 A;Cross-references: UNIPROT:Q03238; GB:L05175; NID:g206691; PIDN:AAA42056.1; PID:g206692  
 C;Superfamily: trypsin; trypsin; trypsin homology  
 C;Keywords: hydrolase; serine proteinase  
 F;21-245/Domain: trypsin homology <TRY>

Query Match 76.6%; Score 36; DB 2; Length 258;  
 Best Local Similarity 66.7%; Pred. No. 29;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
 ||:||||:  
 Db 49 VLVHQRWVL 57

## RESULT 13

D42519  
 A20R protein - vaccinia virus (strain Copenhagen)

C;Species: vaccinia virus

A;Note: host Homo sapiens (man)

C;Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004

C;Accession: D42519

R;Johnson, G.P.

submitted to GenBank, June 1990

A;Reference number: A33172

A;Accession: D42519

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-426 <JOH>

A;Cross-references: UNIPROT:P20995

C;Superfamily: vaccinia virus probable 49.1K protein

Query Match 76.6%; Score 36; DB 2; Length 426;  
 Best Local Similarity 55.6%; Pred. No. 50;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
 ||:||||:  
 Db 130 VLLNNRWIM 138

## RESULT 14

T37408  
 Probable 49.1K protein - vaccinia virus (strain Ankara)

C;Species: vaccinia virus

A;Variety: strain Ankara

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C;Accession: T37408

R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.

submitted to the EMBL Data Library, March 1997

A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strai

A;Reference number: Z20877

A;Accession: T37408

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-426 <ANT>

A;Cross-references: UNIPROT:O57228; EMBL:U94848; PIDN:AAB96523.1

A;Experimental source: strain Ankara

C;Genetics:

A;Note: MVAL32R

C;Superfamily: vaccinia virus probable 49.1K protein

Query Match 76.6%; Score 36; DB 2; Length 426;  
 Best Local Similarity 55.6%; Pred. No. 50;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
 ||:||||:  
 Db 130 VLLNNRWIM 138

## RESULT 15

C72166  
A22R protein - variola minor virus (strain Garcia-1966)  
C:Species: variola minor virus  
C:Date: 24-Nov-1999 #sequence\_revision 24-Nov-1999 #text\_change 09-Jul-2004  
C:Accession: C72166  
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopar  
submitted to GenBank, March 1998  
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor  
A:Reference number: A72150  
A:Accession: C72166  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-426 <SHC>  
A:Cross-references: UNIPROT:Q89166; GB:Y16780; NID:G5830555; PIDN:CAB54724.1; PID:G58306  
A:Experimental source: strain Garcia-1966  
C:Genetics:  
A:Gene: A22R  
C:Superfamily: vaccinia virus probable 49.1K protein

Query Match 76.6%; Score 36; DB 2; Length 426;  
Best Local Similarity 55.6%; Pred. No. 50;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
|||  
DB 130 VLLNNRWIM 138

Search completed: July 13, 2005, 17:31:29  
Job time : 14.5556 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-34

Perfect score: 47

Sequence: 1 VLNERWVL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_prot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	47	100.0	253	1	KLK7_HUMAN	P49862 homo sapien
2	47	100.0	253	2	Q8NSN9	Q8NSN9 homo sapien
3	40	85.1	455	2	Q9FFR7	Q9FFR7 arabidopsis
4	39	83.0	100	2	Q9MYV8	Q9MYV8 bos taurus
5	39	83.0	249	2	Q8JFQ7	Q8JFQ7 gadus morhu
6	38	80.9	223	2	Q9XY58	Q9XY58 ctienocephal
7	38	80.9	234	2	Q7SZE2	Q7SZE2 agkistrodon
8	38	80.9	235	2	Q6EPG8	Q6EPG8 figu rubrip
9	38	80.9	246	1	KLK_PIG	P00752 sus scrofa
10	38	80.9	260	1	NRPN_MOUSE	Q61955 mus musculu
11	38	80.9	260	1	NRPN_RAT	Q88780 rattus norv
12	38	80.9	262	2	Q7Z0G5	Q7Z0G5 phlebotomus
13	38	80.9	313	2	Q8T3B8	Q8T3B8 caenorhabdi
14	38	80.9	377	2	Q9SNY8	Q9SNY8 solanum tub
15	38	80.9	418	2	Q9SNY9	Q9SNY9 solanum tub
16	38	80.9	461	2	P91123	P91123 caenorhabdi
17	38	80.9	791	2	Q7NUH8	Q7NUH8 chromobacte
18	37	78.7	184	2	Q67PW1	Q67PW1 symbiobacte
19	37	78.7	233	2	Q9PT51	Q9PT51 agkistrodon
20	37	78.7	234	2	Q8UUJ1	Q8UUJ1 agkistrodon
21	37	78.7	243	2	Q9VEM7	Q9VEM7 drosophila
22	37	78.7	249	2	Q92046	Q92046 dissostichu
23	37	78.7	249	2	Q788V0	Q788V0 dissostichu
24	37	78.7	249	2	Q9W6K0	Q9W6K0 nototheria
25	37	78.7	250	1	TRIP_PLEPL	P35034 pleuronecte
26	37	78.7	250	2	Q93265	Q93265 pseudopleur
27	37	78.7	254	2	Q97098	Q97098 anopheles a
28	37	78.7	255	2	Q97100	Q97100 anopheles d
29	37	78.7	257	1	VSPA_TRIMU	Q91510 trimeresuru
30	37	78.7	257	1	VSP5_TRIMU	Q91511 trimeresuru
31	37	78.7	257	1	VSP7_TRIMU	Q94584 trimeresuru

32	37	78.7	257	1	VSPC_TRIGA	O13062 trimeresuru
33	37	78.7	257	2	Q71QI5	Q71QI5 trimeresuru
34	37	78.7	257	2	Q71QI7	Q71QI7 trimeresuru
35	37	78.7	257	2	Q71QJ0	Q71QJ0 trimeresuru
36	37	78.7	257	2	Q9PTL3	Q9PTL3 agkistrodon
37	37	78.7	257	2	Q9YVG9	Q9YVG9 agkistrodon
38	37	78.7	257	2	CYR2_ANOGA	Q17025 anopheles g
39	37	78.7	258	2	Q9XY53	Q9XY53 ctienocephal
40	37	78.7	258	2	Q7SZE1	Q7SZE1 gloydius sa
41	37	78.7	259	1	CTR1_ANOGA	Q27289 anopheles g
42	37	78.7	259	2	Q8UUR2	Q8UUR2 crotalus ad
43	37	78.7	260	1	VSP1_TRIJE	Q94668 trimeresuru
44	37	78.7	260	1	VSP2_TRIFL	O13057 trimeresuru
45	37	78.7	262	2	Q8QHK3	Q8QHK3 crotalus at

#### ALIGNMENTS

##### RESULT 1

KLK7\_HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (HSCCE).  
GN Name=KLK7; Synonyms=PRSS6, SCCE;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
RC TISSUE=Skin;  
RX MEDLINE=94308225; PubMed=8034709;  
RA Hanson L., Stromqvist M., Baekman A., Wallbrandt P., Carlstein A., Egelrud T.;  
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";  
RL J. Biol. Chem. 269:19420-19426(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Keratocytes;  
RA Yousef G.M., Scorilas A., Diamandis E.P.;  
RT "Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepel B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Hanson L., Baekman A., Ny A., Edlund M., Edholm E., Tornell J., Wallbrandt P., Egelrud T.;  
RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in mice; a model for chronic itchy dermatitis.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE=95314630; PubMed=7794273;  
RA Skjott A., Stromqvist M., Egelrud T.;  
RT "Primary substrate specificity of recombinant human stratum corneum chymotryptic enzyme.";  
RL Biochem. Biophys. Res. Commun. 211:586-589(1995).  
CC !- FUNCTION: May catalyze the degradation of intercellular cohesive structures in the cornified layer of the skin in the continuous

CC shedding of cells from the skin surface. Specific for amino acid  
CC residues with aromatic side chains in the P1 position. SCCE  
CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
CC phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
CC activation of precursors to inflammatory cytokines.  
CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
CC expressed by keratinocytes in the epidermis. Very low levels are  
CC also seen in the brain and kidney.  
CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L33404; AAC37551.1; -.  
CC EMBL; AF166330; AAD49718.1; -.  
CC EMBL; AF243527; AAG33360.1; -.  
CC EMBL; AF332583; AAK69624.1; -.  
CC PIR; A53968; A53968.  
CC HSSP; P00760; 1EZX.  
CC MEROPS; S01.300; -.  
CC Genew; HGNC:6368; KLK7.  
CC H-InvDB; HIX0015373; -.  
CC MIM; 604438; -.  
CC GO; GO:0008236; F:serine-type peptidase activity; TAS.  
CC GO; GO:0009544; P:epidermal differentiation; TAS.  
CC InterPro; IPR009003; Pept\_Ser\_Cys.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; Trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; Tryp\_SPC; 1.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
KW Signal; Zymogen.  
FT SIGNAL 1 22  
FT PROPEP 23 29 Activation peptide.  
FT CHAIN 30 253 Kallikrein 7.  
FT ACT\_SITE 70 70 Charge relay system (By similarity).  
FT ACT\_SITE 112 112 Charge relay system (By similarity).  
FT ACT\_SITE 205 205 Charge relay system (By similarity).  
FT DISULFID 36 137 By similarity.  
FT DISULFID 55 71 By similarity.  
FT DISULFID 7 239 By similarity.  
FT DISULFID 144 211 By similarity.  
FT DISULFID 176 190 By similarity.  
FT DISULFID 201 226 By similarity.  
FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;

Query Match 100.0%; Score 47; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
DB 58 VLVNERWVL 66  
|||||

RESULT 2  
Q8NSN9 PRELIMINARY; PRT; 253 AA.  
AC Q8NSN9  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2004 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Stratum corneum chymotryptic enzyme, preproprotein.  
GN Name=KLK7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC EMBL; BC032005; AAK32005.1; -.  
CC HSSP; P00760; 1EZX.  
CC GO; GO:0004263; F:chymotrypsin activity; IEA.  
CC GO; GO:0008233; F:peptidase activity; IEA.  
CC GO; GO:0004295; F:trypsin activity; IEA.  
CC GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC InterPro; IPR009003; Pept\_Ser\_Cys.  
CC Pfam; PF00089; Trypsin; 1\_Ser\_Cys.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; Tryp\_SPC; 1.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 100.0%; Score 47; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
DB 58 VLVNERWVL 66  
|||||

RESULT 3  
Q9FFR7 PRELIMINARY; PRT; 455 AA.  
AC Q9FFR7  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:MBK23.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97471969; PubMed=9330910;  
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence  
 RT features of the 1.6 Mb regions covered by twenty physically assigned  
 RT P1 clones.";  
 RL DNA Res. 4:215-230(1997).  
 DR EMBL; AB005233; BAB11469.1; --  
 DR InterPro; IPR001810; F-box.  
 DR Pfam; PF00646; F-box; 1.  
 SQ SEQUENCE 455 AA; 52305 MW; 763F5D5278D23F13 CRC64;

Query Match 85.1%; Score 40; DB 2; Length 455;  
 Best Local Similarity 75.0%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LVNERNVL 9  
 Db :|||||  
 104 IINERNVL 111

## RESULT 4

O9MYV8  
 ID Q9MYV8 PRELIMINARY; PRT; 100 AA.  
 AC Q9MYV8;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Haptooglobin (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oviduct;  
 RA Lavery K.S., Gabler C., Killian G.J.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ271156; CAC00531.1; --  
 DR HSP; P00751; IDLE.  
 DR MEROPS; S01.972; --  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 1  
 FT CHAIN 100 100  
 SQ SEQUENCE 100 AA; 11239 MW; FFF6AE5AB43CED9 CRC64;

Query Match 83.0%; Score 39; DB 2; Length 100;  
 Best Local Similarity 75.0%; Pred. No. 12;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LVNERNVL 9  
 Db :|||||  
 22 LINERNVL 29

## RESULT 5

O9JF07  
 ID Q9JF07 PRELIMINARY; PRT; 249 AA.  
 AC Q9JF07;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Trypsinogen Y precursor (Fragment).  
 GN Name=tryp gene;  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pyloric caecum;  
 RX PubMed=10612685;  
 RA Spilliaert R.;  
 RT "Atlantic Cod Trypsin Y - Member of a Novel Trypsin Group.";  
 RL Mar. Biotechnol. 1:598-607(1999).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Pyloric caecum;  
 RA Gudmundsdottir A.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; AJ459311; CAD30563.1; --  
 DR HSP; P00746; 1FDP.  
 DR MEROPS; S01.124; --

DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease; Signal.  
 FT SIGNAL 1 15 Potential.  
 FT CHAIN 23 >249 trypsinogen Y.  
 FT NON\_TER 23 249  
 SQ SEQUENCE 249 AA; 27400 MW; 8FB98462CEDBEFC9 CRC64;

Query Match 83.0%; Score 39; DB 2; Length 249;  
 Best Local Similarity 66.7%; Pred. No. 31;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLVNERNVL 9  
 Db |||:|||||  
 50 VLINDQWVL 58

## RESULT 6

O9XY58  
 ID Q9XY58 PRELIMINARY; PRT; 223 AA.  
 AC Q9XY58;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Chymotrypsin-like serine protease (Fragment).  
 GN Name=SP-4;  
 OS Ctenocephalides felis (Cat flea).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Siphonaptera; Pulicidae; Pulicinae;  
 OC Ctenocephalides.  
 OX NCBI\_TaxID=7515;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gaines P.J., Sampson C.M., Rushlow K.E., Stiegler G.L.;  
 RT "Cloning of a family of serine protease genes from the cat flea  
 RT Ctenocephalides felis.";  
 RL Insect Mol. Biol. 8:11-22(1999).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.

```
DR EMBL; AF053916; AAD21836.1; -.
DR HSSP; P00761; 1EPT.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR Pfam; PF00089; Trypsin; 1_Ser_Cys.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 223 AA; 24844 MW; C34EBE9455DD7DE9 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 223;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVNERWVL 9
Db :|||||
35 ILNERWVL 42

RESULT 7
Q7SZE2 PRELIMINARY; PRT; 234 AA.
AC Q7SZE2
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Thrombin-like enzyme defibrase (fragment).
OS Agkistrodon caliginosus (Korean viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Crotalinae; Gloydius.
OX NCBI_TaxID=35671;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Snake venom gland;
RA Sun D.-J., Yang T.-S.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY204242; AAP20637.1; -.
DR HSSP; Q91516; 1BQY.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 234 AA; 25351 MW; 0124C1485525F560 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 234;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVNERWVL 9
Db :|||||
30 LVNEEWVL 37

RESULT 8
Q66PG8 PRELIMINARY; PRT; 235 AA.
AC Q66PG8
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Trypsinogen (fragment).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC Roach J.C.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY661446; AAU06121.1; -.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1_Ser_Cys.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 235 AA; 26025 MW; 489D3617C943CCSA CRC64;

Query Match 80.9%; Score 38; DB 2; Length 235;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLNERWVL 9
Db :|||||
36 VLINKQWVL 44

RESULT 9
KLK_PIG STANDARD; PRT; 246 AA.
AC P00752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Glandular kallikrein precursor (EC 3.4.21.35) (Tissue kallikrein).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-15 AND 95-102.
RX MEDLINE=69230643; PubMed=3246040;
RA Kanada M., Aoki K., Ikeita M., Kizuki K., Moriya H., Kamo M.,
RA Tsugita A.;
RT "Generation of alpha- and beta-kallikreins from porcine pancreatic
RT prokallikrein by the action of trypsin.";
RL Chem. Pharm. Bull. 36:4891-4899(1988).
RN [2]
RP SEQUENCE OF 8-87 AND 95-246.
RC TISSUE=Pancreas;
RA Tschesche H., Mair G., Godec G., Fiedler F., Ehret W., Hirschauer C.,
RA Lemon M., Fritz H., Schmidt-Kastner G., Kutzbach C.;
RT "The primary structure of porcine glandular kallikreins.";
RL Adv. Exp. Med. Biol. 120:245-260(1979).
RN [3]
RP SEQUENCE OF 8-87; 95-127 AND 176-246.
RC TISSUE=Pancreas;
RA Ehret W.;
RT "The primary structure of the kallikrein from porcine pancreas.";
RL Thesis (1976), University of Munich, Germany.
RN [4]
```

SEQUENCE OF 84-98.  
 RX MEDLINE=90336040; PubMed=2379280;  
 RA Kanada M., Ikekita M., Kurahashi T., Aoki K., Kizuki K., Moriya H.,  
 RA Sweeley C.C., Kano M., Tsugita A.;  
 RT "Generation of a different type of beta-kallikrein from porcine  
 RT pancreatic alpha-kallikrein by the action of chymotrypsin --  
 RT observation of proteolytic processing occurring around 'kallikrein  
 RT autolysis loop' region.";  
 RL Chem. Pharm. Bull. 38:1053-1057(1990).  
 RN [5]  
 RN SEQUENCE OF 128-175.  
 RC TISSUE=Pancreas;  
 RA Ehret W.;  
 RT "Investigation of the sequence of amino acid residues 127 to 174 of  
 RT the kallikrein from porcine pancreas.";  
 RL Theiss (1978), University of Munich, Germany.  
 RN [6]  
 RN REVIEW.  
 RX MEDLINE=82194877; PubMed=7043199;  
 RA Fiedler F., Fink E., Tschesche H., Fritz H.;  
 RT "Porcine glandular kallikreins.";  
 RL Meth. Enzymol. 80:493-532(1981).  
 RN [7]  
 RX X-RAY CRYSTALLOGRAPHY (2 ANGSTROMS), AND REVISIONS.  
 RA Bode W., Chen Z., Bartels K., Kutzbach C., Schmidt-Kastner G.,  
 RA Bartunik H.;  
 RT "Refined 2-A X-ray crystal structure of porcine pancreatic kallikrein  
 RT A, a specific trypsin-like serine proteinase. Crystallization,  
 RT structure determination, crystallographic refinement, structure and  
 RT its comparison with bovine trypsin.";  
 RL J. Mol. Biol. 164:237-282(1983).  
 RN [8]  
 RX X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF COMPLEX WITH BOVINE  
 RP PANCREATIC TRYPSIN INHIBITOR.  
 RX MEDLINE=83189107; PubMed=6551452;  
 RA Chen Z., Bode W.;  
 RT "Refined 2.5 A X-ray crystal structure of the complex formed by  
 RT porcine kallikrein A and the bovine pancreatic trypsin inhibitor.  
 RT Crystallization, Patterson search, structure determination,  
 RT refinement, structure and comparison with its components and with the  
 RT bovine trypsin-pancreatic trypsin inhibitor complex.";  
 RL J. Mol. Biol. 164:283-311(1983).  
 RN [9]  
 RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH HIRUSTASIN.  
 RX MEDLINE=97184690; PubMed=9032072; DOI=10.1016/S0969-2126(97)00183-4;  
 RA Mittl P.R.E., di Marco S., Fendrich G., Pohlig G., Heim J.,  
 RA Sommerhoff C., Fritz H., Priestle J.P., Gruetter M.G.;  
 RT "A new structural class of serine protease inhibitors revealed by the  
 RT structure of the hirustasin-kallikrein complex.";  
 RL Structure 5:253-264(1997).  
 RN [10]  
 RP ERRATUM.  
 RA Mittl P.R.E., di Marco S., Fendrich G., Pohlig G., Heim J.,  
 RA Sommerhoff C., Fritz H., Priestle J.P., Gruetter M.G.;  
 RL Structure 5:585-585(1997).  
 RN [11]  
 RX STRUCTURE OF CARBOHYDRATES.  
 RA MEDLINE=89062455; PubMed=3196708;  
 RA Tomiya N., Yamaguchi T., Awaya J., Kurono M., Endo S., Arata Y.,  
 RA Takahashi N., Ishihara H., Mori M., Tejima S.;  
 RT "Structural analyses of asparagine-linked oligosaccharides of porcine  
 RT pancreatic kallikrein.";  
 RL Biochemistry 27:7146-7154(1988).  
 CC -!- FUNCTION: Glandular kallikreins cleave Met-Lys and Arg-Ser bonds  
 CC in kininogen to release Lys-bradykinin.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage of Arg-|-Xaa bonds in  
 CC small molecule substrates. Highly selective action to release  
 CC kallidin (lysyl-bradykinin) from kininogen involves hydrolysis of  
 CC Met-|-Xaa or Leu-|-Xaa.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
 CC subfamily.

-!- CAUTION: Native porcine kallikrein is a monomer. Chains of the  
 CC pancreatic beta-kallikrein are heterogeneous artifacts of the  
 CC proteolytic degradation during isolation.  
 DR PDB; 1HIA; X-ray; A/X=8-87, B/Y=95-246.  
 DR PDB; 2KAI; X-ray; A=8-87, B=95-246.  
 DR PDB; 2PKA; X-ray; A/X=8-87, B/Y=95-246.  
 DR MEROPS; S01.160; --  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW 3D-structure; Direct protein sequencing; Glycoprotein; Hydrolase;  
 KW Pancreas; Serine protease; Zymogen.  
 FT PROPEP 1 7  
 FT CHAIN 8 246 Glandular kallikrein.  
 FT DOMAIN 85 104 Kallikrein (autolysis) loop.  
 FT DISULFID 14 158  
 FT DISULFID 33 49  
 FT DISULFID 135 204  
 FT DISULFID 169 183  
 FT DISULFID 194 219  
 FT CARBOHYD 85 85  
 FT CARBOHYD 239 239  
 FT ACT\_SITE 48 48  
 FT ACT\_SITE 103 103  
 FT ACT\_SITE 198 198  
 FT STRAND 9 9  
 FT STRAND 12 13  
 FT TURN 16 17  
 FT TURN 20 21  
 FT STRAND 22 27  
 FT TURN 28 29  
 FT STRAND 30 39  
 FT TURN 40 41  
 FT STRAND 42 45  
 FT HELIX 47 49  
 FT STRAND 55 58  
 FT STRAND 62 62  
 FT TURN 63 64  
 FT STRAND 71 80  
 FT TURN 82 86  
 FT TURN 101 102  
 FT STRAND 105 109  
 FT STRAND 116 116  
 FT TURN 117 118  
 FT STRAND 119 119  
 FT TURN 131 132  
 FT STRAND 134 139  
 FT STRAND 155 155  
 FT STRAND 157 164  
 FT TURN 166 166  
 FT TURN 167 170  
 FT TURN 171 172  
 FT TURN 179 180  
 FT STRAND 181 185  
 FT TURN 187 188  
 FT STRAND 192 192  
 FT TURN 195 196  
 FT TURN 198 199  
 FT STRAND 201 204  
 FT TURN 205 206  
 FT STRAND 207 214  
 FT TURN 219 220  
 FT TURN 222 223  
 FT STRAND 226 230  
 FT TURN 231 234  
 FT HELIX 235 244  
 SQ SEQUENCE 246 AA; 27170 MW; 5991CEDE406A19A1 CRC64;

Query Match 80.9%; Score 38; DB 1; Length 246;  
 Best Local Similarity 77.8%; Pred. No. 48;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWYL 9  
 DB 36 VLVNPKWYL 44  
 ||||| :|||  
 ||||| :|||

RESULT 10  
 NRPN MOUSE  
 ID NRPN MOUSE STANDARD; PRT; 260 AA.  
 AC Q61955;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Neuropilin precursor (BC 3.4.21.-) (NP) (Kallikrein 8).  
 GN Name=Klk8; Synonyms=Nrp1, Prs19;  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Hippocampus;  
 RX MEDLINE=95348817; PubMed=7623137;  
 RA Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,  
 RA Ito J., Nishino H., Aimoto S., Kiyama H., Shiosaka S.,  
 RT "Expression and activity-dependent changes of a novel limbic-serine  
 RT protease gene in the hippocampus.",  
 RL J. Neurosci. 15:5088-5097 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yoshida S., Hirata A., Inoue N., Shiosaka S.;  
 RT "Cloning and assignment of mouse neuropilin gene, Prs19 to chromosome  
 RT 7B4.",  
 RN [3]  
 RP Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Czech II; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.F., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF N-TERMINUS, CHARACTERIZATION, AND MASS SPECTROMETRY.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=98225202; PubMed=9556608; DOI=10.1074/jbc.273.18.11189;  
 RA Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,  
 RA Shiosaka T., Midorikawa R., Kamachi T., Kawabe A., Shiosaka S.;  
 RT "Characterization of recombinant and brain neuropilin, a plasticity-  
 RT related serine protease.",  
 RL J. Biol. Chem. 273:11189-11196 (1998).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 33-257.

RC TISSUE=Hippocampus;  
 RX MEDLINE=95134351; PubMed=9933620; DOI=10.1074/jbc.274.7.4220;  
 RA Kishi T., Kato M., Shimizu T., Kato K., Matsumoto K., Yoshida S.,  
 RA Shiosaka S., Hakoshima T.;  
 RT "Crystal structure of neuropilin, a hippocampal protease involved in  
 RT kindling epileptogenesis.",  
 RL J. Biol. Chem. 274:4220-4224 (1999).  
 CC -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and  
 CC hippocampal plasticity. Has a strong proteolytic activity against  
 CC fibronectin.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.  
 CC -!- ENZYME REGULATION: Strongly inhibited by diisopropyl  
 CC fluorophosphate, leupeptin and (4-aminophenyl)methanesulfonyl 1-  
 CC fluoride.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed specifically in the limbic system of  
 CC mouse brain and is localized at highest concentration in pyramidal  
 CC neurons of the hippocampal CA1-3 subfields.  
 CC -!- MASS SPECTROMETRY: MW=26613; METHOD=MALDI; RANGE=29-260;  
 CC NOTE=Ref. 4.  
 CC -!- MASS SPECTROMETRY: MW=26229; METHOD=MALDI; RANGE=33-260;  
 CC NOTE=Ref. 4.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
 CC subfamily.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; D30785; BAA06451.1; --  
 EMBL; AB032202; BAA92435.1; --  
 EMBL; BC055895; AAH55895.1; --  
 PIR; I56559; I56559.  
 PDB; INPM; X-ray; A/B=33-256.  
 MEROPS; S01-244; --  
 MGD; MGI:892018; Klk8.  
 InterPro; IPR009003; Pept\_Ser\_Cys.  
 InterPro; IPR01254; Peptidase\_S1.  
 InterPro; IPR001314; Peptidase\_S1A.  
 Pfam; PF00089; Trypsin; 1.  
 PRINTS; PR00722; CHYMOTRYPSIN.  
 PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 PROSITE; PS00135; TRYPSIN\_SER; 1.  
 3D-structure; Direct protein sequencing; Glycoprotein; Hydrolase;  
 serine protease; Signal; Zymogen.  
 SIGNAL 1 28 Potential.  
 PROPEP 29 32  
 CHAIN 33 260 Neuropilin.  
 ACT SITE 73 73 Charge relay system (By similarity).  
 ACT SITE 120 120 Charge relay system (By similarity).  
 ACT SITE 212 212 Charge relay system (By similarity).  
 DISULFID 39 173  
 DISULFID 58 74  
 DISULFID 145 246  
 DISULFID 152 218  
 DISULFID 184 198  
 DISULFID 208 233  
 CARBOHYD 110 110 N-linked (GlcNAc...) (Potential).  
 STRAND 34 34  
 TURN 37 38  
 TURN 41 42  
 TURN 45 46  
 STRAND 47 52  
 TURN 53 54  
 STRAND 55 64  
 TURN 65 66  
 TURN 67 70  
 HELIX 72 74

FT STRAND 80 83  
 FT TURN 87 87  
 FT TURN 88 89  
 FT STRAND 96 98  
 FT TURN 100 105  
 FT TURN 107 108  
 FT TURN 114 115  
 FT TURN 118 119  
 FT STRAND 122 126  
 FT STRAND 140 141  
 FT TURN 148 149  
 FT STRAND 151 156  
 FT STRAND 170 170  
 FT STRAND 172 178  
 FT HELIX 181 187  
 FT TURN 189 191  
 FT TURN 194 195  
 FT STRAND 196 200  
 FT TURN 202 203  
 FT STRAND 206 206  
 FT TURN 209 210  
 FT TURN 212 213  
 FT STRAND 215 218  
 FT TURN 219 220  
 FT STRAND 221 228  
 FT STRAND 235 235  
 FT TURN 236 237  
 FT STRAND 238 238  
 FT STRAND 240 244  
 FT HELIX 245 256  
 SQ SEQUENCE 260 AA; BESF6P6BE37CD60E CRC64;

Query Match 80.9%; Score 38; DB 1; Length 260;  
 Best Local Similarity 77.8%; Pred. No. 51;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
 ||| :|||  
 Db 61 VLVGDRWVL 69

RESULT 11  
 NRPN RAT STANDARD; PRT; 260 AA.  
 AC O88780;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Neuropilin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Brain serine protease 1).  
 GN Name=Klk8; Synonyms=Bsp1, Nrpn, Prss19;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=Fischer; TISSUE=Brain;  
 RA MEDLINE=90389725; PubMed=9722524; DOI=10.1074/jbc.273.36.23004;  
 RX Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathe R.;  
 FT "Serine proteases in rodent hippocampus."  
 RL J. Biol. Chem. 273:23004-23011(1998).  
 CC -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and hippocampal plasticity. Has a strong proteolytic activity against fibronectin (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.  
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- TISSUE SPECIFICITY: Restricted to hippocampus.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein subfamily.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AJ005641; CAA06643.1; -  
 DR HSSP; Q61955; INPM.  
 DR MEROPS; S01.244; -  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Hydrolase; Serine protease; Signal; Zymogen.  
 FT SIGNAL 1 28 Potential.  
 FT PROPEP 29 32 By similarity.  
 FT CHAIN 33 260 Neuropilin.  
 FT ACT\_SITE 73 73 Charge relay system (By similarity).  
 FT ACT\_SITE 120 120 Charge relay system (By similarity).  
 FT ACT\_SITE 212 212 Charge relay system (By similarity).  
 FT DISULFID 39 173 By similarity.  
 FT DISULFID 58 74 By similarity.  
 FT DISULFID 145 246 By similarity.  
 FT DISULFID 152 218 By similarity.  
 FT DISULFID 184 198 By similarity.  
 FT DISULFID 208 233 By similarity.  
 FT CARBOHYD 110 110 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 260 AA; 58DF4F0602A0B7F5 CRC64;

Query Match 80.9%; Score 38; DB 1; Length 260;  
 Best Local Similarity 77.8%; Pred. No. 51;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
 ||| :|||  
 Db 61 VLVGDRWVL 69

RESULT 12  
 Q720G5 PRELIMINARY; PRT; 262 AA.  
 ID Q720G5;  
 AC Q720G5;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Chymotrypsin.  
 GN Name=chym1;  
 OS Phlebotomus papatasi.  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Psychodoidea;  
 OC Psychodidae; Phlebotomus.  
 OX NCBI\_TaxID=29031;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC TISSUE=Midgut;  
 RA Ramalho-Ortigao M., Kanhawi S., Rowton E., Sacks D.L., Ribeiro J.M.,  
 RA Valenzuela J.G.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY128106; AAM96938.1; -  
 DR HSSP; P18291; 1P18.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.

```
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 262 AA; 28863 MW; 31F11012206A65D4 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 262;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVNERWVL 9
Db 57 ILNERWVL 64
:::|||||

RESULT 13
Q8T3B8 PRELIMINARY; PRT; 313 AA.
AC Q8T3B8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C32E8.6.
GN Name=C32E8.6; ORFNames=C32E8.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome Sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S.;
RT "The sequence of C. elegans cosmid C32E8.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
family.
DR EMBL; U88308; AAM15557.1; -.
DR HSRP; P08659; ILCI.
DR WormBase; WBGene00016324; C32E8.6.
DR WormPep; C32E8.6b; CE30371.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 35608 MW; 4372ED66478E434A CRC64;

Query Match 80.9%; Score 38; DB 2; Length 313;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERW 7
Db 219 VLVNERW 225
:::|||||

RESULT 14
Q9SNY8 PRELIMINARY; PRT; 377 AA.
ID Q9SNY8
AC Q9SNY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Branched-chain amino acid aminotransferase (EC 2.6.1.42).
GN Name=BCAT2;
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RA Campbell M.A., Patel J.K., Meyers J.L., Myrick L.C., Gustin J.L.;
RT "Genes encoding for branched-chain amino acid aminotransferase are
differentially expressed in plants.";
RL Plant Physiol. Biochem. 39:855-860(2001).
DR EMBL; AF193846; AAF07192.1; -.
DR HSRP; P00510; IILK.
DR GO; GO:0004084; F: branched-chain-amino-acid transaminase acti. .; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0009081; P: branched chain family amino acid metabolism; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR001544; Aminotrans_IV.
DR InterPro; IPR005786; B_aminotransII.
DR Pfam; PF01063; Aminotran_4; 1.
DR ProDom; PD001961; Aminotrans_IV; 1.
DR TIGFAMS; TIGR01123; ilvE II; 1.
DR PROSITE; PS00770; AA_TRANSFER_CLASS_4; 1.
KW Aminotransferase; Transferase.
SQ SEQUENCE 377 AA; 40924 MW; 8CC8460A5F15B646 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 377;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLVNERWV 8
Db 147 VLVNERWI 154
:::|||||

RESULT 15
Q9SNY9 PRELIMINARY; PRT; 418 AA.
ID Q9SNY9
AC Q9SNY9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Branched-chain amino acid aminotransferase (EC 2.6.1.42).
GN Name=BCAT1;
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RA Campbell M.A., Patel J.K., Meyers J.L., Myrick L.C., Gustin J.L.;
RT "Genes encoding for branched-chain amino acid aminotransferase are
differentially expressed in plants.";
RL Plant Physiol. Biochem. 39:855-860(2001).
DR EMBL; AF193845; AAF07191.1; -.
DR HSRP; O15382; 1KT8.
```

DR GO; GO:0004084; F:branched-chain-amino-acid transaminase acti. . .; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0009081; P:branched chain family amino acid metabolism; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR001544; AminoTrans IV.  
DR InterPro; IPR005786; B\_amino\_transII.  
DR Pfam; PF01063; AminoTran\_4; 1.  
DR ProDom; PD001961; AminoTrans IV; 1.  
DR TIGRFAMs; TIGR01123; ilvE\_II; 1.  
KW AminoTransferase; Transferase.  
SQ SEQUENCE 418 AA; 45295 MW; E6EA6C2B3A58C814 CRC64;

Query Match 80.9%; Score 38; DB 2; Length 418;  
Best Local Similarity 75.0%; Pred.No. 84;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVNERWV 8  
|||  
Db 188 VLANERWI 195

Search completed: July 13, 2005; 17:29:22  
Job time : 64.2222 secs

**This page blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-34

Perfect score: 47

Sequence: 1 VLVNERWVL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	9	3	US-09-502-600-34
2	47	100.0	9	4	US-09-918-243-34
3	47	100.0	154	3	US-09-261-416-7
4	47	100.0	224	3	US-08-944-483-33
5	47	100.0	225	2	US-08-557-146-12
6	47	100.0	225	2	US-09-027-337-4
7	47	100.0	225	2	US-09-154-344-12
8	47	100.0	225	4	US-09-644-600-4
9	47	100.0	225	4	US-09-654-600A-4
10	47	100.0	253	2	US-08-557-146-2
11	47	100.0	253	2	US-08-824-874-3
12	47	100.0	253	2	US-09-154-344-2
13	47	100.0	253	3	US-08-930-188-2
14	47	100.0	253	3	US-09-210-084-3
15	47	100.0	253	4	US-09-764-762-3
16	47	100.0	253	5	PCT-US96-04294-2
17	47	100.0	265	4	US-09-949-016-7716
18	43	91.5	9	3	US-09-502-600-40
19	43	91.5	9	3	US-09-502-600-50
20	43	91.5	9	4	US-09-918-243-40
21	43	91.5	9	4	US-09-918-243-50
22	39	83.0	228	3	US-08-944-483-44
23	39	83.0	253	6	5223425-8
24	39	83.0	253	6	5223425-8
25	38	80.9	64	1	US-08-485-455D-75
26	38	80.9	64	2	US-08-482-130C-75
27	38	80.9	64	2	US-08-484-211C-75

```

28 38 80.9 64 3 US-08-906-769-75 Sequence 75, Appl
29 38 80.9 64 3 US-08-906-616-75 Sequence 75, Appl
30 38 80.9 64 3 US-08-817-795-75 Sequence 75, Appl
31 38 80.9 64 3 US-08-485-443B-75 Sequence 75, Appl
32 38 80.9 64 3 US-08-639-075A-75 Sequence 75, Appl
33 38 80.9 64 3 US-09-012-431-75 Sequence 75, Appl
34 38 80.9 64 3 US-09-012-692-75 Sequence 75, Appl
35 38 80.9 64 3 US-08-906-613-75 Sequence 75, Appl
36 38 80.9 64 5 PCT-US95-14442A-75 Sequence 75, Appl
37 38 80.9 223 1 US-08-485-455D-17 Sequence 17, Appl
38 38 80.9 223 2 US-08-482-130C-17 Sequence 17, Appl
39 38 80.9 223 2 US-08-484-211C-17 Sequence 17, Appl
40 38 80.9 223 3 US-08-817-795-17 Sequence 17, Appl
41 38 80.9 223 3 US-08-485-443B-17 Sequence 17, Appl
42 38 80.9 223 5 PCT-US95-14442A-17 Sequence 17, Appl
43 38 80.9 224 3 US-08-906-769-17 Sequence 17, Appl
44 38 80.9 224 3 US-08-906-616-17 Sequence 17, Appl
45 38 80.9 224 3 US-08-639-075A-17 Sequence 17, Appl

```

#### ALIGNMENTS

```

RESULT 1
US-09-502-600-34
; Sequence 34, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 34
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 58-66 of the SCCE protein
US-09-502-600-34

```

```

Query Match 100.0%; Score 47; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 VLVNERWVL 9
Db 1 VLVNERWVL 9

```

```

RESULT 2
US-09-918-243-34
; Sequence 34, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 34
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```
;
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 58-66 of the SCCE protein
US-09-918-243-34

Query Match      100.0%; Score 47; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9
Db 1 VLVNERWVL 9

RESULT 3
US-09-261-416-7
; Sequence 7, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261.416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 7
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
; OTHER INFORMATION: chymotryptic enzyme (Scce) homologous to similar
; OTHER INFORMATION: domain in TADG-12
US-09-261-416-7

Query Match      100.0%; Score 47; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9
Db 4 VLVNERWVL 12

RESULT 4
US-08-944-483-33
; Sequence 33, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

;
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183-US-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-33

Query Match      100.0%; Score 47; DB 3; Length 224;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9
Db 29 VLVNERWVL 37

RESULT 5
US-08-557-146-12
; Sequence 12, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-557-146-12

Query Match      100.0%; Score 47; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLVNERWVL 9
DB      30 VLVNERWVL 38

RESULT 6
US-09-027-337-4
; Sequence 4, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
; FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027,337B
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to
; OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4

Query Match      100.0%; Score 47; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLVNERWVL 9
DB      30 VLVNERWVL 38

RESULT 7
US-09-154-344-12
; Sequence 12, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
```

```
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-154-344-12

Query Match      100.0%; Score 47; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLVNERWVL 9
DB      30 VLVNERWVL 38

RESULT 8
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-644-600-4

Query Match      100.0%; Score 47; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLVNERWVL 9
DB      30 VLVNERWVL 38

RESULT 9
US-09-654-600A-4
; Sequence 4, Application US/09654600A
; Patent No. 6649741
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/654,600A
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
```

09/027,337  
 ; PRIOR FILING DATE: 1999-10-20  
 ; 1998-02-20  
 ; NUMBER OF SEQ ID NOS: 98  
 ; SEQ ID NO 4  
 ; LENGTH: 225  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: SCCE  
 US-09-654-600A-4

Query Match 100.0%; Score 47; DB 4; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 0.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
 Db 30 VLVNERWVL 38  
 |||||

RESULT 10  
 US-08-557-146-2  
 ; Sequence 2, Application US/08557146  
 ; Patent No. 5834290  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Egelrud, Torbjorn  
 ; APPLICANT: Hansson, Lennart  
 ; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
 ; TITLE OF INVENTION: Enzyme (SCCE)  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: White & Case, Patent Department  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036-2787  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/557,146  
 ; FILING DATE: 14-DEC-1995  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Sterner, Richard J.  
 ; REGISTRATION NUMBER: 35,372  
 ; REFERENCE/DOCKET NUMBER: 1103326-181  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 819-8783  
 ; TELEFAX: (212) 354-8113  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 253 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-557-146-2

Query Match 100.0%; Score 47; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.68;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
 Db 58 VLVNERWVL 66  
 |||||

RESULT 11  
 US-08-824-874-3

; Sequence 3, Application US/08824874  
 ; Patent No. 5962300  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Lal, Preeti  
 ; TITLE OF INVENTION: NOVEL KALLIKREIN  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/824,874  
 ; FILING DATE: Filed Herewith  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0252 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 253 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 532504  
 US-08-824-874-3

Query Match 100.0%; Score 47; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.68;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
 Db 58 VLVNERWVL 66  
 |||||

RESULT 12  
 US-09-154-344-2  
 ; Sequence 2, Application US/09154344  
 ; Patent No. 5981256  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Egelrud, Torbjorn  
 ; APPLICANT: Hansson, Lennart  
 ; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
 ; TITLE OF INVENTION: Enzyme (SCCE)  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: White & Case, Patent Department  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036-2787  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,344  
FILING DATE: 16-SEP-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-154-344-2

Query Match 100.0%; Score 47; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 VLVNERWVL 9  
DB 58 VLVNERWVL 66

RESULT 13  
US-09-930-188-2  
Sequence 2, Application US/08930188  
Patent No. 6093397  
GENERAL INFORMATION:  
APPLICANT: Dixon, Eric P.  
APPLICANT: Johnstone, Edward M.  
APPLICANT: Little, Shella P.  
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
TITLE OF INVENTION: RELATED NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930,188  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/416,257  
FILING DATE: 04-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Blalock, Donna K.  
REGISTRATION NUMBER: 38,082  
REFERENCE/DOCKET NUMBER: X9239  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-277-1090  
TELEFAX: 317-276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-930-188-2

Query Match 100.0%; Score 47; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 VLVNERWVL 9  
DB 58 VLVNERWVL 66

RESULT 14  
US-09-210-084-3  
Sequence 3, Application US/09210084  
Patent No. 6197511  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-SEQ for Windows Version 2.0  
CURRENT APPLICATION DATA: US/09/210,084  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,874  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0252 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 532504  
US-09-210-084-3

Query Match 100.0%; Score 47; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.68; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 VLVNERWVL 9  
DB 58 VLVNERWVL 66

RESULT 15  
US-09-764-762-3

; Sequence 3, Application US/09764762  
; Patent No. 6472195  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09764,762  
; FILING DATE: 16-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/210,084  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 532504  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-764-762-3

Query Match 100.0%; Score 47; DB 4; Length 253;  
Best Local Similarity 100.0%; Pred.No. 0.68;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
| | | | |  
Db 58 VLVNERWVL 66

Search completed: July 13, 2005, 17:34:26  
Job time : 19.4444 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds  
(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-34

Perfect score: 47

Sequence: 1 VLVNERWVL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match.100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pap.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pap.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pap.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pap.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pap.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pap.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pap.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pap.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pap.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pap.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pap.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	9	US-09-918-243-34	Sequence 34, Appl
2	47	100.0	9	US-09-905-083-34	Sequence 34, Appl
3	47	100.0	9	US-10-372-521-34	Sequence 34, Appl
4	47	100.0	9	US-10-831-075-34	Sequence 100, Appl
5	47	100.0	97	US-10-262-511-100	Sequence 100, Appl
6	47	100.0	181	US-10-262-511-98	Sequence 98, Appl
7	47	100.0	198	US-10-262-511-96	Sequence 96, Appl
8	47	100.0	224	US-09-789-210-33	Sequence 33, Appl
9	47	100.0	224	US-10-262-511-104	Sequence 104, Appl
10	47	100.0	225	US-10-600-187-4	Sequence 4, Appl
11	47	100.0	247	US-10-262-511-102	Sequence 102, Appl

12	47	100.0	250	15	US-10-262-511-92	Sequence 92, Appl
13	47	100.0	252	15	US-10-262-511-94	Sequence 94, Appl
14	47	100.0	253	9	US-09-888-615-98	Sequence 98, Appl
15	47	100.0	253	9	US-09-764-762-3	Sequence 3, Appl
16	47	100.0	253	14	US-10-071-214-2	Sequence 2, Appl
17	47	100.0	253	14	US-10-071-214-48	Sequence 48, Appl
18	47	100.0	253	14	US-10-264-283-90	Sequence 90, Appl
19	47	100.0	253	15	US-10-235-027-498	Sequence 498, Appl
20	47	100.0	253	15	US-10-173-999-48	Sequence 48, Appl
21	47	100.0	253	16	US-10-408-765A-639	Sequence 639, Appl
22	47	100.0	253	16	US-10-643-795A-95	Sequence 95, Appl
23	47	100.0	253	17	US-10-948-518-95	Sequence 95, Appl
24	47	100.0	253	17	US-10-868-490A-1	Sequence 1, Appl
25	47	100.0	257	15	US-10-344-394-38	Sequence 38, Appl
26	43	91.5	9	9	US-09-918-243-40	Sequence 40, Appl
27	43	91.5	9	9	US-09-918-243-50	Sequence 50, Appl
28	43	91.5	9	9	US-09-905-083-40	Sequence 40, Appl
29	43	91.5	9	9	US-09-905-083-50	Sequence 50, Appl
30	43	91.5	9	15	US-10-372-521-40	Sequence 40, Appl
31	43	91.5	9	15	US-10-372-521-50	Sequence 50, Appl
32	43	91.5	9	16	US-10-831-075-40	Sequence 40, Appl
33	43	91.5	9	16	US-10-831-075-50	Sequence 50, Appl
34	40	85.1	243	14	US-10-071-214-46	Sequence 46, Appl
35	40	85.1	249	14	US-10-071-214-47	Sequence 47, Appl
36	39	83.0	228	11	US-09-789-210-44	Sequence 44, Appl
37	39	83.0	228	14	US-10-202-676-6	Sequence 6, Appl
38	38	80.9	81	13	US-10-045-676A-7	Sequence 7, Appl
39	38	80.9	81	14	US-10-170-789-57	Sequence 57, Appl
40	38	80.9	225	15	US-10-072-012-611	Sequence 611, Appl
41	38	80.9	259	15	US-10-072-012-222	Sequence 222, Appl
42	38	80.9	260	9	US-09-796-294-8	Sequence 8, Appl
43	38	80.9	260	14	US-10-461-787-8	Sequence 8, Appl
44	38	80.9	260	15	US-10-072-012-609	Sequence 609, Appl
45	38	80.9	260	15	US-10-072-012-610	Sequence 610, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-918-243-34  
; Sequence 34, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 34  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 58-66 of the SCCE protein  
US-09-918-243-34

Query Match 100.0%; Score 47; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9

Db 1 VLVNERWVL 9

##### RESULT 2

US-09-905-083-34  
; Sequence 34, Application US/09905083  
; Patent No. US20020146708A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; TITLE OF INVENTION: Ovarian Cancer  
; FILE REFERENCE: D6223CIP/C/Div  
; CURRENT APPLICATION NUMBER: US/09/905,083  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 09/502,600  
; PRIOR FILING DATE: 2000-02-11  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 34  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 58-66 of the SCCE protein  
US-09-905-083-34

Query Match 100.0%; Score 47; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
Db 1 VLVNERWVL 9  
|||||

RESULT 3  
US-10-372-521-34  
; Sequence 34, Application US/10372521  
; Publication No. US20030223973A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP2  
; CURRENT APPLICATION NUMBER: US/10/372,521  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 34  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 58-66 of the SCCE protein  
US-10-372-521-34

Query Match 100.0%; Score 47; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
Db 1 VLVNERWVL 9  
|||||

RESULT 4  
US-10-831-075-34  
; Sequence 34, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 34  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 58-66 of the SCCE protein  
US-10-831-075-34

Query Match 100.0%; Score 47; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
Db 1 VLVNERWVL 9  
|||||

RESULT 5  
US-10-262-511-100  
; Sequence 100, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennnda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038



; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/373,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 100  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-100

Query Match 100.0%; Score 47; DB 15; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.8; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1 VLVNERWVL 9  
Db 39 VLVNERWVL 47

## RESULT 6

US-10-262-511-98  
; Sequence 98, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 98  
; LENGTH: 181  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-98

Query Match 100.0%; Score 47; DB 15; Length 181;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
Db 39 VLVNERWVL 47

## RESULT 7

US-10-262-511-96  
; Sequence 96, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917

;  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 96  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-262-511-96

Query Match 100.0%; Score 47; DB 15; Length 198;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
Db 58 VLVNERWVL 66

RESULT 8

US-09-789-210-33  
; Sequence 33, Application US/09789210  
; Publication No. US20040241646A1

GENERAL INFORMATION:

; APPLICANT: COHEN, MAURICE  
; COLPITS, TRACEY L.  
; FRIEDMAN, PAULA N.  
; GRANADOS, EDWARD N.  
; KLASS, MICHAEL R.  
; RUSSELL, JOHN C.  
; STEWART, KENT D.  
; STROUPE, STEVEN D.

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; OF THE PROSTATE

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500

COMPUTER READABLE FORM:

; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/789,210  
; FILING DATE: 20-FEB-2001  
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/944,483  
; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 224 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: None  
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
; US-09-789-210-33

Query Match 100.0%; Score 47; DB 11; Length 224;

Best Local Similarity 100.0%; Pred. No. 6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVNERWVL 9  
Db 29 VLVNERWVL 37

RESULT 9

US-10-262-511-104

; Sequence 104, Application US/10262511  
; Publication No. US20040038223A1

GENERAL INFORMATION:

; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Beighs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-462C

CURRENT APPLICATION NUMBER: US/10/262,511

CURRENT FILING DATE: 2003-05-28

PRIOR APPLICATION NUMBER: 60/326,483

PRIOR FILING DATE: 2001-10-02

PRIOR APPLICATION NUMBER: 60/373,815

PRIOR FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: 60/327,917

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/381,642

PRIOR FILING DATE: 2002-05-17

PRIOR APPLICATION NUMBER: 60/328,029

PRIOR FILING DATE: 2002-10-09

PRIOR APPLICATION NUMBER: 60/381,038

PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/373,260  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; PRIOR FILING DATE: 2001-10-05  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 104  
 ; LENGTH: 224  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-104

Query Match 100.0%; Score 47; DB 15; Length 224;

Best Local Similarity 100.0%; Pred. No. 6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
Db 44 VLVNERWVL 52

## RESULT 10

US-10-600-187-4  
 ; Sequence 4, Application US/10600187  
 ; Publication No. US20040086910A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, Timothy J.  
 ; APPLICANT: Tanimoto, Hirotoishi  
 ; TITLE OF INVENTION: TAG-15: An Extracellular Serine Protease  
 ; TITLE OF INVENTION: Overexpressed in Carcinomas  
 ; FILE REFERENCE: D6064CJP/D  
 ; CURRENT APPLICATION NUMBER: US/10/600,187  
 ; CURRENT FILING DATE: 2003-06-20  
 ; PRIOR APPLICATION NUMBER: US/09/654,600A  
 ; PRIOR FILING DATE: 2000-09-01  
 ; PRIOR APPLICATION NUMBER: 09/421,213  
 ; 09/027,337  
 ; PRIOR FILING DATE: 1999-10-20  
 ; 1998-02-20  
 ; NUMBER OF SEQ ID NOS: 98  
 ; SEQ ID NO 4  
 ; LENGTH: 225  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: SCCE  
 US-10-600-187-4

Query Match 100.0%; Score 47; DB 15; Length 225;

Best Local Similarity 100.0%; Pred. No. 6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
Db 30 VLVNERWVL 38

## RESULT 11

US-10-262-511-102  
 ; Sequence 102, Application US/10262511  
 ; Publication No. US20040038223A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glenda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Ju, Jingfang  
 ; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Saaha)  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Malvankar, Uriel M.  
 ; APPLICANT: Ort, Tatiana  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Zernusen, Bryan D.  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Pena, Carol E. A.  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Shimkets, Richard A.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Berghs, Constance  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-462C  
 ; CURRENT APPLICATION NUMBER: US/10/262,511  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/373,815  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,642  
 ; PRIOR FILING DATE: 2002-05-17  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,038  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/373,260  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; PRIOR FILING DATE: 2001-10-05  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 102  
 ; LENGTH: 247  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-102

Query Match 100.0%; Score 47; DB 15; Length 247;

Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVNERWVL 9  
Db 44 VLVNERWVL 52

## RESULT 12

US-10-262-511-92  
 ; Sequence 92, Application US/10262511  
 ; Publication No. US20040038223A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glenda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.

```
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Stone, David J.
; APPLICANT: Rastelli, Luca
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agée, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-03-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 92
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-92
```

```
Query Match 100.0%; Score 47; DB 15; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VLVNERWVL 9
Db 55 VLVNERWVL 63
```

```
RESULT 13
US-10-262-511-94
; Sequence 94, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agée, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 94
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-94
```

```
Query Match 100.0%; Score 47; DB 15; Length 252;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VLVNERWVL 9
Db 57 VLVNERWVL 65
```

```
RESULT 14
US-09-888-615-98
```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083a-35

Perfect score: 40

Sequence: 1 LLPLQLILL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	AAE08240	Human str
2	40	100.0	9	ADR68796	Human str
3	40	100.0	136	ABG23378	Novel hum
4	40	100.0	198	ADA05736	Human NOV
5	40	100.0	198	ADN62900	Human NOV
6	40	100.0	250	ADA05732	Human NOV
7	40	100.0	250	ADN62896	Human NOV
8	40	100.0	253	AAR67888	Human str
9	40	100.0	253	AAW05383	Human any
10	40	100.0	253	ABB84421	Human SCC
11	40	100.0	253	ABB84406	Human SCC
12	40	100.0	253	AAU82740	Amino aci
13	40	100.0	253	ABU07440	Protein d
14	40	100.0	253	ABU07471	Protein d
15	40	100.0	253	ABR58471	Human str
16	40	100.0	253	ADB80484	Ovarian c
17	40	100.0	253	ADJ68833	Human hea
18	40	100.0	253	ADN39180	Cancer/an
19	40	100.0	253	ADJ06515	Human tum
20	40	100.0	253	ADN04182	Antipsori
21	40	100.0	253	ADR72880	Human ova
22	40	100.0	257	AAB21326	Human HSC
23	38	95.0	142	AAB63580	Human gas
24	38	95.0	156	AAB63578	Human gas
25	38	95.0	159	AAB63582	Human gas

26	36	90.0	9	AAE08238	Human str
27	36	90.0	9	ADR68794	Human str
28	36	90.0	845	ABU23951	Protein e
29	36	90.0	868	ABD21222	Bacterial
30	35	87.5	23	ABB43858	Peptide #
31	35	87.5	23	AAAM3771	Peptide #
32	35	87.5	23	AAAM64837	Human bra
33	35	87.5	23	ABG59233	Human liv
34	35	87.5	23	ABG46617	Human pep
35	35	87.5	480	AAW98431	H. pylori
36	34	85.0	70	AAAY36291	Human sec
37	34	85.0	70	ADAL1692	Human nov
38	34	85.0	73	AAW89030	Polypepti
39	34	85.0	73	ABBS1130	Human sec
40	34	85.0	73	ABO45387	Novel hum
41	34	85.0	73	ABO26867	Protein a
42	34	85.0	86	AAO08397	Human pol
43	34	85.0	86	ABO00884	Polypepti
44	34	85.0	89	ADE09276	Novel pro
45	34	85.0	89	ADS12009	Human the

## ALIGNMENTS

## RESULT 1

AAE08240  
ID AAE08240 standard; peptide; 9 AA.

XX AAE08240;

XX 01-NOV-2001 (first entry)

XX Human stratum corneum chymotrypsin enzyme peptide #5 (residues 6-14).

XX Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;  
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;  
KW antisense therapy; malignant hyperplasia.

XX Homo sapiens.

XX WO200159158-A1.

XX 16-AUG-2001.

XX 07-FEB-2001; 2001WO-US003977.

XX 11-FEB-2000; 2000US-00502600.

XX (UYAR-) UNIV ARKANSAS.

XX O'brien TJ;

XX WPI; 2001-514676/56.

XX Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX Claim 25; Page 103; 127pp; English.

XX The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer. CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

SQ Sequence 9 AA;

Query Match 100.0%; Score 40; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 1 LLPLQILL 9

RESULT 2  
ADRG68796  
ID ADRG68796 standard; peptide; 9 AA.  
XX AC  
XX ADRG68796;  
XX DT  
XX 02-DEC-2004 (first entry)  
XX DE Human stratum corneum chymotryptic enzyme peptide fragment SEQ ID NO:35.  
XX KW serine protease; stratum corneum chymotryptic enzyme; SCCE;  
XX KW immune response; ovarian cancer; lung cancer; prostate cancer;  
XX KW pancreatic cancer; colon cancer.  
XX OS Homo sapiens.  
XX PN WO2004075723-A2.  
XX PD 10-SEP-2004.  
XX PF 20-FEB-2004; 2004WO-US005134.  
XX PR 21-FEB-2003; 2003US-00372521.  
XX PA (UYAR-) UNIV ARKANSAS.  
XX PI O'brien TJ, Cannon MJ, Santin A;  
XX DR WPI; 2004-653294/63.  
XX PT Use of stratum corneum chymotryptic enzyme (SCCE) peptides, for  
PT vaccinating an individual against SCCE, and in monitoring and developing  
PT immunotherapies for ovarian and other malignancies.  
XX PS Claim 5; SEQ ID NO 35; 117pp; English.  
XX CC The invention relates to a novel method for vaccinating an individual  
CC against stratum corneum chymotryptic enzyme (SCCE), comprising inoculating  
CC an individual with a SCCE peptide, which elicits an immune response in  
CC the individual. A peptide of the invention acts as a stratum corneum  
CC chymotryptic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
CC an individual against SCCE, particularly an individual having, suspected  
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
CC an individual, such as ovarian, breast, lung, colon, prostate, or  
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
CC peptides are also useful in the monitoring and development of  
CC immunotherapies for ovarian and other malignancies. The present sequence  
CC represents a peptide fragment of serine protease SCCE (stratum corneum  
CC chymotryptic enzyme).  
XX SQ Sequence 9 AA;  
Query Match 100.0%; Score 40; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 1 LLPLQILL 9

RESULT 3  
ABG23378  
ID ABG23378 standard; protein; 136 AA.

XX AC ABG23378;  
XX DT 18-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #23369.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX DR WPI; 2001-639362/73.  
XX DR N-PSDB; AAS87565.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX PS Claim 20; SEQ ID NO 53737; 103pp; English.  
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 136 AA;  
Query Match 100.0%; Score 40; DB 4; Length 136;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 12 LLPLQILL 20

RESULT 4  
ADA05736  
ID ADA05736 standard; protein; 198 AA.  
XX AC ADA05736;



XX 06-NOV-2003 (first entry)  
 XX Human NOV18c protein SEQ ID NO:96.  
 DE human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX Homo sapiens.  
 XX WO2003029424-A2.  
 XX 10-APR-2003.  
 XX 02-OCT-2002; 2002WO-US031373.  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 01-NOV-2001; 2001US-0349575P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX (CURA-) CURAGEN CORP.  
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX WPI: 2003-381626/36.  
 DR N-PSDB; ADA05735.  
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX Claim 1; Page 170; 586pp; English.  
 PS The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.  
 XX Sequence 198 AA;  
 SQ  
 Query Match 100.0%; Score 40; DB 6; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 LLPLQILL 9  
 Db 6 LLPLQILL 14  
 RESULT 5  
 ID ADN62900 standard; protein; 198 AA.  
 XX AC ADN62900;  
 XX DT 01-JUL-2004 (first entry)  
 XX DE Human NOV18c.  
 XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX Homo sapiens.  
 XX US2004038223-A1.  
 XX 26-FEB-2004.  
 XX 01-OCT-2002; 2002US-00262511.  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339268P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JULJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATI/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RASI/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIFI/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytke KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2004-213931/20.  
 DR N-PSDB; ADN62899.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 96; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 198 AA;

Query Match 100.0%; Score 40; DB 8; Length 198;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
 |||||  
 Db 6 LLPLQILL 14

RESULT 6

ADA05732  
 ID ADA05732 standard; protein; 250 AA.

XX ADA05732;

XX 06-NOV-2003 (first entry)

XX Human NOV18a protein SEQ ID NO:92.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipaemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0345357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383658P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Smithson G., Millett I., Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Pattarajan M, Sytek KA, Edinger SE, Ellerman K, Malyankar UM;  
 PI Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 DR WPI; 2003-381626/36.  
 DR N-PSDB; ADA05731.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 169-170; 586pp; English.  
 XX  
 CC The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.  
 XX  
 SQ Sequence 250 AA;  
 Query Match 100.0%; Score 40; DB 6; Length 250;  
 Best Local Similarity 100.0%; Pred. NO. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQILL 9  
 Db 3 LLPLQILL 11  
 RESULT 7  
 ADN62896  
 ID ADN62896 standard; protein; 250 AA.  
 XX  
 AC ADN62896;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human NOV18a.  
 XX  
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2004038223-A1.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 01-OCT-2002; 2002US-00262511.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 29-OCT-2001; 2001US-0343629P.  
 PR 01-NOV-2001; 2001US-0345357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383658P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX  
 PA (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.  
PA (JUUJ/) JU J.  
PA (LILL/) LI L.  
PA (GUOX/) GUO X.  
PA (PATT/) PATTURAJAN M.  
PA (SPYT/) SPYTEK K A.  
PA (EDIN/) EDINGER S R.  
PA (ELLE/) ELLERMAN K.  
PA (MALY/) MALYANKAR U M.  
PA (ORTT/) ORT T.  
PA (GORM/) GORMAN L.  
PA (ZERH/) ZERHUSEN B D.  
PA (ANDE/) ANDERSON D W.  
PA (ZHON/) ZHONG M.  
PA (CATT/) CATTERTON E.  
PA (JIWW/) JI W.  
PA (MILL/) MILLER C E.  
PA (RAST/) RASTELLI L.  
PA (STON/) STONE D J.  
PA (PENA/) PENNA C E A.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.  
PA (ROTH/) ROTHENBERG M E.  
PA (LEAC/) LEACH M D.  
PA (AGEE/) AGEE M L.  
PA (BERG/) BERGHS C.  
PA (DIPI/) DIPIPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI: 2004-213931/20.  
DR N-PSDB: ADN62895.  
XX  
XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
PS Claim 1; SEQ ID NO 92; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression or  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 250 AA;  
Query Match 100.0%; Score 40; DB 8; Length 250;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLPLQILL 9  
Db 3 LLPLQILL 11  
RESULT 8  
AAR67888  
ID AAR67888 standard; protein; 253 AA.  
XX  
AC AAR67888;  
XX  
DT 25-MAR-2003 (revised)  
DT 09-AUG-1995 (first entry)  
XX  
DE Human stratum corneum chymotrophic recombinant enzyme (SCCE).  
XX  
KW Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
KW callosities; keratosis pilaris; ichthyoses; eczema.  
XX  
OS Homo sapiens.  
XX  
PN WO9500651-A1.  
XX  
PD 05-JAN-1995.  
XX  
PF 20-JUN-1994; 94WO-IB000166.  
XX  
PR 18-JUN-1993; 93DK-00000725.  
XX  
PA (SYMB-) SYMBICOM AB.  
XX  
PI Egelrud T, Hansson L;  
XX  
DR WPI: 1995-052088/07.  
DR N-PSDB: AAQ81203.  
XX  
PT Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
PT related vectors, transformed cells and polypeptides, useful for treating  
PT skin disorders, e.g. acne or psoriasis, and for identification of  
PT specific inhibitors.  
XX  
PS Disclosure; Page 97; 137pp; English.  
XX  
CC The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
CC and skin care products, especially to treat and prevent acne, xeroderma,  
CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
CC recombinantly following mammal, insect, plant, or microorganism  
CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX  
SQ Sequence 253 AA;  
Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLPLQILL 9  
Db 6 LLPLQILL 14  
RESULT 9  
AAW05383

ID AAW05383 standard; protein; 253 AA.  
AC AAW05383;  
XX  
DT 31-DEC-1996 (first entry)  
XX  
DE Human amyloid precursor protein protease.  
XX  
KW Amyloid precursor protein protease; Alzheimer's disease; diagnosis;  
KW therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO9631122-A1.  
XX  
PD 10-OCT-1996.  
XX  
PF 02-APR-1996; 96WO-US004294.  
XX  
PR 04-APR-1995; 95US-00416257.  
XX  
PA (ELIL ) LILLY & CO ELI.  
XX  
PI Dixon EP, Johnstone EM, Little SP;  
XX  
DR WPI; 1996-464694/46.  
DR N-PSDB; AAT39783.  
XX  
PT New isolated human amyloid precursor protein protease - used to develop  
PT prods. for the treatment or diagnosis of associated conditions, esp.  
PT Alzheimer's disease.  
XX  
PS Claim 1; Page 44-45; 55pp; English.  
XX  
CC Human amyloid precursor protein protease (AAW05383) is involved in the  
CC processing or clearance of amyloid precursor protein to form beta-amyloid  
CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)  
CC obtd. from a human lung library. Recombinant protease can be produced in  
CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic  
CC (partic. AV-120 host cells. It is used to develop products for the design  
CC and testing of cpds. useful for treating or preventing conditions  
CC associated with beta-amyloid peptide, esp. Alzheimer's disease  
XX  
SQ Sequence 253 AA;  
Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLPLQILL 9  
DB 6 LLPLQILL 14  
RESULT 10  
ABB84421  
ID ABB84421 standard; peptide; 253 AA.  
XX  
AC ABB84421;  
XX  
DT 08-NOV-2002 (first entry)  
XX  
DE Human SCCE protein N-terminal fragment SQ ID 48.  
XX  
KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.  
XX  
OS Homo sapiens.  
XX  
PN WO200262135-A2.  
XX

PD 15-AUG-2002.  
XX  
PF 08-FEB-2002; 2002WO-IB001300.  
XX  
PR 09-FEB-2001; 2001CA-02332655.  
PR 09-FEB-2001; 2001DK-00000218.  
XX  
PA (EGEL/) EGELRUD T.  
PA (HANS/) HANSSON L.  
XX  
PI Egelrud T, Hansson L;  
XX  
DR WPI; 2002-643380/69.  
XX  
PT Transgenic mammal or its embryo useful as model for human disease, has  
PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
PT enzyme operably linked to promoter that drives its expression in skin.  
XX  
PS Example 6; Page 37; 74pp; English.  
XX  
CC This invention describes a novel non-human transgenic mammal or mammalian  
CC embryo having integrated within its genome, a heterologous nucleotide  
CC sequence comprising at least a significant part of a nucleotide sequence  
CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant.  
CC operably linked to a promoter that drives expression of heterologous scce  
CC or its variant in skin. The product of the invention is useful as a model  
CC for the study of disease with the aim of improving treatment, to relieve  
CC or ameliorate a pathological condition, for development or testing of a  
CC cosmetic or a pharmaceutical formulation, and for the development of a  
CC diagnostic method. It can also be used as a model for a skin disease or  
CC skin cancer. The invention is also useful for screening or identifying a  
CC compound or composition effective for the prevention or treatment of an  
CC abnormal or unwanted phenotype, and for screening or identifying a  
CC compound or composition effective for the prevention or treatment of  
CC inflammatory skin diseases selected from diseases consisting of epidermal  
CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
CC as a model for further studies of itch mechanisms and the testing of  
CC potential compounds and compositions for relieve of various skin diseases  
CC where itch is a component. This sequence represents the N-terminal  
CC fragment of the human stratum corneum chymotryptic enzyme, SCCE  
CC synonymous with human kallikrein 7 (KLK7), used in the development of the  
CC transgenic mammals described in the invention  
XX  
SQ Sequence 253 AA;  
Query Match 100.0%; Score 40; DB 5; Length 253;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLPLQILL 9  
DB 6 LLPLQILL 14  
RESULT 11  
ABB84406  
ID ABB84406 standard; protein; 253 AA.  
XX  
AC ABB84406;  
XX  
DT 08-NOV-2002 (first entry)  
XX  
DE Human SCCE protein.  
XX  
KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.  
XX  
OS Homo sapiens.  
XX

PN WO2000262135-A2.  
 XX 15-AUG-2002.  
 XX 08-FEB-2002; 2002WO-IB001300.  
 XX 09-FEB-2001; 2001CA-02332655.  
 PR 09-FEB-2001; 2001DK-00000218.  
 XX (EGEL/) EGELRUD T.  
 PA (HANS/) HANSSON L.  
 XX Egelrud T, Hansson L;  
 XX WPI; 2002-643380/69.  
 DR N-PSDB; ABQ76226.  
 XX Transgenic mammal or its embryo useful as model for human disease, has  
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
 PT enzyme operably linked to promoter that drives its expression in skin.  
 XX  
 PS Claim 10; Page 58-59; 74pp; English.  
 XX This invention describes a novel non-human transgenic mammal or mammalian  
 CC embryo having integrated within its genome, a heterologous nucleotide  
 CC sequence comprising at least a significant part of a nucleotide sequence  
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
 CC operably linked to a promoter that drives expression of heterologous scce  
 CC or its variant in skin. The product of the invention is useful as a model  
 CC for the study of disease with the aim of improving treatment, to relieve  
 CC or ameliorate a pathogenic condition, for development or testing of a  
 CC cosmetic or a pharmaceutical formulation, and for the development of a  
 CC diagnostic method. It can also be used as a model for a skin disease or  
 CC skin cancer. The invention is also useful for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of an  
 CC abnormal or unwanted phenotype, and for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of  
 CC inflammatory skin diseases selected from diseases consisting of epidermal  
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
 CC as a model for further studies of itch mechanisms and the testing of  
 CC potential compounds and compositions for relieve of various skin diseases  
 CC where itch is a component. This sequence represents the human stratum  
 CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymous  
 CC with human kallikrein 7 (KLK7) and is used in the development of the  
 CC transgenic mammals described in the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQLILL 9  
 Db 6 LLPLQLILL 14  
 RESULT 12  
 AAU82740  
 ID AAU82740 standard; protein; 253 AA.  
 XX  
 AC AAU82740;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE Amino acid sequence of novel human protease #39.  
 XX  
 KW Human; protease; cancer; immune-related disorder; cardiovascular disease;  
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;

KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
 KW ocular disease; cytostatic; enzyme.  
 OS Homo sapiens.  
 XX WO200200860-A2.  
 XX 03-JAN-2002.  
 XX 26-JUN-2001; 2001WO-US020171.  
 XX 26-JUN-2000; 2000US-0214047P.  
 PR (SUGE-) SUGEN INC.  
 PA Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
 PI Charyczak G;  
 XX WPI; 2002-139913/18.  
 DR N-PSDB; ABK31782.  
 XX Nucleic acids encoding novel human proteases, useful for useful for  
 PT treating diseases and disorders such as cancers, immune-related diseases  
 PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory  
 PT disorders.  
 PS Claim 6; Fig 2N; 313pp; English.  
 XX The present invention relates to the isolation of novel human proteases,  
 CC and the nucleic acids encoding them. The sequences of the invention are  
 CC useful for treating diseases and disorders such as cancers (e.g. breast,  
 CC colon, lung), immune-related diseases and disorders (e.g. inflammatory  
 CC diseases and asthma), cardiovascular diseases (e.g. restenosis and  
 CC coronary thrombosis), brain or neuronal-associated diseases, metabolic  
 CC disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.  
 CC rheumatoid arthritis and psoriasis), central or peripheral nervous system  
 CC diseases, migraines, pain, sexual dysfunction, mood disorders, attention  
 CC disorders, cognitive disorders, hypotension, hypertension, psychotic  
 CC disorders, neurological disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease) and dyskinesias. The nucleic acids and polypeptides are also  
 CC useful for treating viral infections caused by human immunodeficiency  
 CC virus (HIV), and non-viral infections such as ocular disease (e.g.  
 CC glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel  
 CC human proteases of the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQLILL 9  
 Db 6 LLPLQLILL 14  
 RESULT 13  
 ABU07440  
 ID ABU07440 standard; protein; 253 AA.  
 XX  
 AC ABU07440;  
 XX  
 DT 28-JAN-2003 (first entry)  
 XX  
 DE Protein differentially regulated in prostate cancer #43.  
 XX  
 KW Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX  
 OS Homo sapiens.  
 XX WO200281638-A2.

XX PD 17-OCT-2002.  
 XX FF 08-APR-2002; 2002WO-US010824.  
 XX PR 06-APR-2001; 2001US-0281731P.  
 XX PR 06-APR-2001; 2001US-0281732P.  
 XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 XX PI Sun Z, Jay G;  
 XX DR WPI; 2003-058520/05.  
 XX DR N-PSDB; ABX10343.  
 XX Novel genes which are differentially regulated in prostate cancer, useful  
 PT for diagnosing prostate cancer in prostate tissue sample and assessing  
 PT therapeutic or preventive intervention in prostate cancer patients.  
 XX  
 PS Claim 1; Page 293-294; 416pp; English.  
 XX The invention describes genes (I) which are differentially regulated in  
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 CC sample comprising prostate tissue, which involves determining the number  
 CC of target genes which are differentially-regulated in the sample, where  
 CC the number is indicative of the probability that the sample comprises  
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 CC intervention in a subject having a prostate cancer, which involves  
 CC determining the expression levels in a sample comprising prostate tissue  
 CC of target genes which are differentially-regulated in prostate cancer.  
 CC Preferably, the expression levels of at least 10 genes are determined.  
 CC (I) is also useful for identifying agents that modulate a biological  
 CC activity of a polypeptide differentially-regulated in prostate cancer  
 CC cells, which involves contacting a polypeptide differentially-regulated  
 CC in prostate cancer cells with a test agent under conditions effective for  
 CC the test agent to modulate a biological activity of the polypeptide, and  
 CC determining whether the test agent modulates the biological activity. (I)  
 CC is useful as molecular markers, as drug targets, and for detecting,  
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
 CC preventing or treating, determining predisposition to diseases and  
 CC conditions especially relating to prostate cancer. (I) and its expression  
 CC products are used in the diagnostic test to assay for presence of cancer  
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 CC of cancer, its stage of development, the nature of genetic defect, etc.  
 CC The polypeptide encoded by (I) can be used as target for therapy or drug  
 CC discovery. (I) can also be used for expressing the polypeptide and thus  
 CC for searching specific binding partners of the polypeptide. (I) is useful  
 CC in therapeutic applications to treat prostate cancer. The identification  
 CC of specific genes, and groups of genes, expressed in pathways  
 CC physiologically relevant to prostate cancer permits the definition of  
 CC functional and disease pathways and the delineation of targets in these  
 CC pathways which are useful in diagnostic, therapeutic, and clinical  
 CC applications. This is the amino acid sequence of a protein differentially  
 CC regulated in prostate cancer  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLPLQLLL 9  
 Db |||||  
 6 LLPLQLLL 14  
 RESULT 14  
 ABU07471  
 ID ABU07471 standard; protein; 253 AA.  
 XX  
 AC ABU07471;  
 XX

DT 28-JAN-2003 (first entry)  
 DE Protein differentially regulated in prostate cancer #74.  
 XX Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX Homo sapiens.  
 XX WO200281638-A2.  
 XX PN 17-OCT-2002.  
 XX PD 08-APR-2002; 2002WO-US010824.  
 XX PF 06-APR-2001; 2001US-0281731P.  
 XX PR 06-APR-2001; 2001US-0281732P.  
 XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 XX PI Sun Z, Jay G;  
 XX DR WPI; 2003-058520/05.  
 XX DR N-PSDB; ABX10375.  
 XX Novel genes which are differentially regulated in prostate cancer, useful  
 PT for diagnosing prostate cancer in prostate tissue sample and assessing  
 PT therapeutic or preventive intervention in prostate cancer patients.  
 XX  
 PS Claim 1; Page 351; 416pp; English.  
 XX The invention describes genes (I) which are differentially regulated in  
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 CC sample comprising prostate tissue, which involves determining the number  
 CC of target genes which are differentially-regulated in the sample, where  
 CC the number is indicative of the probability that the sample comprises  
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 CC intervention in a subject having a prostate cancer, which involves  
 CC determining the expression levels in a sample comprising prostate tissue  
 CC of target genes which are differentially-regulated in prostate cancer.  
 CC Preferably, the expression levels of at least 10 genes are determined.  
 CC (I) is also useful for identifying agents that modulate a biological  
 CC activity of a polypeptide differentially-regulated in prostate cancer  
 CC cells, which involves contacting a polypeptide differentially-regulated  
 CC in prostate cancer cells with a test agent under conditions effective for  
 CC the test agent to modulate a biological activity of the polypeptide, and  
 CC determining whether the test agent modulates the biological activity. (I)  
 CC is useful as molecular markers, as drug targets, and for detecting,  
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
 CC preventing or treating, determining predisposition to diseases and  
 CC conditions especially relating to prostate cancer. (I) and its expression  
 CC products are used in the diagnostic test to assay for presence of cancer  
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 CC of cancer, its stage of development, the nature of genetic defect, etc.  
 CC The polypeptide encoded by (I) can be used as target for therapy or drug  
 CC discovery. (I) can also be used for expressing the polypeptide and thus  
 CC for searching specific binding partners of the polypeptide. (I) is useful  
 CC in therapeutic applications to treat prostate cancer. The identification  
 CC of specific genes, and groups of genes, expressed in pathways  
 CC physiologically relevant to prostate cancer permits the definition of  
 CC functional and disease pathways and the delineation of targets in these  
 CC pathways which are useful in diagnostic, therapeutic, and clinical  
 CC applications. This is the amino acid sequence of a protein differentially  
 CC regulated in prostate cancer  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
| | | | |  
Db 6 LLPLQILL 14

## RESULT 15

ABR58471  
ID ABR58471 standard; protein; 253 AA.  
XX  
AC ABR58471;  
XX  
DT 07-JUL-2003 (first entry)  
XX  
DE Human stratum corneum chymotryptic enzyme - ovarian cancer clone 01676P.  
XX  
KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029468-A1.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031467.  
XX  
PR 02-OCT-2001; 2001US-0327135P.  
XX  
PR 30-MAY-2002; 2002US-0384531P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
FI Algate PA, Mannion J;  
XX  
XX WPI; 2003-372001/35.  
XX  
PT New polynucleotide and polypeptide useful for diagnosing and/or treating  
PT cancer, particularly ovarian cancer, and as a vaccine.  
XX  
PS Claim 2; Page 157-158; 169pp; English.  
XX

CC The invention relates to a novel isolated polynucleotide. The  
CC polynucleotides of the invention have cytostatic activity, and may have a  
CC use in gene therapy, and in a vaccine. The composition and methods are  
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.  
CC The composition may also be used as a vaccine to prevent cancer. The  
CC present sequence is used in the exemplification of the invention  
XX

SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
| | | | |  
Db 6 LLPLQILL 14

Search completed: July 13, 2005, 17:19:53  
Job time : 77.6667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083A-35

Perfect score: 40

Sequence: 1 LLPLQILL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	253	2	A53968	serine proteinase
2	36	90.0	845	2	D97163	cation transport p
3	35	87.5	493	2	A71875	hypothetical prote
4	35	87.5	496	2	B64638	conserved hypothet
5	34	85.0	266	2	JC7300	tax-responsive ele
6	34	85.0	267	2	JC4857	hepatocarcinogenes
7	34	85.0	369	2	S33603	surfactant protein
8	34	85.0	371	1	JN0450	conglutinin precu
9	34	85.0	371	1	I45878	conglutinin - bovi
10	34	85.0	429	2	T17215	hypothetical prote
11	34	85.0	539	2	H82994	probable permease
12	34	85.0	1092	2	T13520	hypothetical prote
13	34	85.0	1236	2	T50904	Mg protoporphyrin
14	33	82.5	146	2	H75201	hypothetical prote
15	33	82.5	277	2	H84314	cytochrome aa3 con
16	33	82.5	282	2	H70890	hypothetical prote
17	33	82.5	294	2	E82257	cholera toxin tran
18	33	82.5	294	2	A25970	transcription acti
19	33	82.5	653	2	A46362	amyloid precursor-
20	32	80.0	245	2	E75440	hypothetical prote
21	32	80.0	253	2	AF1535	transport protein
22	32	80.0	256	2	JQ0708	lipopolysaccharide
23	32	80.0	394	1	B43750	chloramphenicol re
24	32	80.0	394	2	B85535	probable transport
25	32	80.0	394	2	F90684	probable transport
26	32	80.0	412	2	A57468	P-selectin glycopor
27	32	80.0	491	2	JC6197	stromelysin 3 (EC
28	32	80.0	492	2	A44399	stromelysin 3 (EC
29	32	80.0	543	2	AI2088	Na+/H+-exchanging

30	32	80.0	1041	2	T29010	hypothetical prote
31	32	80.0	1306	1	A31759	peptidyl-dipeptida
32	32	80.0	1485	2	S23756	CFTR protein - Afr
33	31	77.5	157	2	A83214	probable transcript
34	31	77.5	189	2	T43766	hypothetical prote
35	31	77.5	196	2	G65039	hypothetical prote
36	31	77.5	209	2	A54984	ELF-1 protein prec
37	31	77.5	213	2	JE0322	ephraim-A2 - human
38	31	77.5	303	2	C84914	hypothetical prote
39	31	77.5	306	2	H86811	sugar ABC transpor
40	31	77.5	327	1	A41720	acid phosphatase (
41	31	77.5	328	1	A41927	insulin-like growt
42	31	77.5	370	2	A83334	daunorubicin resis
43	31	77.5	372	2	A98157	probable permease
44	31	77.5	372	2	AH3130	ABC transporter, m
45	31	77.5	388	2	T26360	hypothetical prote

ALIGNMENTS

RESULT 1

A53968  
serine proteinase SCCE precursor - human  
N:Alternate names: stratum corneum chymotryptic enzyme  
C:Species: Homo sapiens (man)  
C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: A53968  
R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.  
J. Biol. Chem. 269, 19420-19426, 1994  
A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme  
A:Reference number: A53968; MUID:94308225; PMID:8034709  
A:Accession: A53968  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-253 <HAN>  
A:Cross-references: UNIPROT:P49862; GB:L33404; NID:9521214; PIDN:AAC37551.1; PID:9532504  
C:Genetics:  
A:Map position: 7q35-7q35  
A:Superfamily: trypsin; trypsin homology  
P:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 40; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9

DB 6 LLPLQILL 14

RESULT 2

D97163  
cation transport P-type ATPase CAC2137 [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: D97163  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
.; Daly, M.G.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: D97163  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-845 <KUR>  
A:Cross-references: UNIPROT:Q97H76; GB:AE001437; PIDN:AAK80095.1; PID:gi15025128; GSPDB:G

Query Match	90.0%;	Score 36;	DB 2;	Length 845;
Best Local Similarity	77.8%;	Pred. No. 40;		
Matches	7;	Conservative	2;	Mismatches 0; Indels 0; Gaps 0;
QY	1	LLPLQIILL 9		
DB	670	LLPIQILLI 678		
RESULT 3				
A:1875				
Hypothetical protein jhp0880 - Helicobacter pylori (strain J99)				
C:Species: Helicobacter pylori				
A:Variety: strain J99				
C:Date: 12-Feb-1999	#sequence_revision 12-Feb-1999	#text_change 09-Jul-2004		
C:Accession: A71875				
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999				
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path				
A:Reference number: A71800;	MUID:99120557;	PMID:9923682		
A:Accession: A71875				
A>Status: preliminary				
A:Molecule type: DNA				
A:Residues: 1-493 <ARN>				
A:Cross-references: UNIPROT:Q9ZKQ5; GB:AE001518; GB:AE001439; NID:G4155454; PIDN:AAD0646				
A:Experimental source: strain J99				
C:Genetics:				
A:Gene: jhp0880				
Query Match	87.5%;	Score 35;	DB 2;	Length 493;
Best Local Similarity	88.9%;	Pred. No. 37;		
Matches	8;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	LLPLQIILL 9		
DB	254	LLPLSILL 262		
RESULT 4				
B64638				
conserved hypothetical integral membrane protein HP0946 - Helicobacter pylori (strain 26)				
C:Species: Helicobacter pylori				
C:Date: 09-Aug-1997	#sequence_revision 09-Aug-1997	#text_change 09-Jul-2004		
C:Accession: B64638				
R:Tomb, J.F.; White, O.; Kervilavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliak, H.G.; Glodek, A.; McKenne son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997				
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.				
A:Reference number: A64520;	MUID:97394467;	PMID:9252185		
A:Accession: B64638				
A>Status: preliminary;	nucleic acid sequence not shown;	translation not shown		
A:Molecule type: DNA				
A:Residues: 1-496 <TOM>				
A:Cross-references: UNIPROT:Q25600; GB:AE000604; GB:AE000511; NID:G2314086; PIDN:AAD0799				
C:Genetics:				
A:Start codon: GTG				
Query Match	87.5%;	Score 35;	DB 2;	Length 496;
Best Local Similarity	88.9%;	Pred. No. 37;		
Matches	8;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	LLPLQIILL 9		
DB	254	LLPLSILL 262		
RESULT 5				
JC7300				
tax-responsive element-binding protein 5 - mouse				
C:Species: Mus musculus (house mouse)				

C:Date: 18-Aug-2000	#sequence_revision 18-Aug-2000	#text_change 09-Jul-2004		
C:Accession: JC7300				
R:Masaki, T.; Noguchi, H.; Kobayashi, M.; Yoshida, M.; Takamatsu, K. DNA Res. 7, 187-193, 2000				
A:Title: Isolation and characterization of the gene encoding mouse tax-responsive element				
A:Reference number: JC7300				
A:Accession: JC7300				
A:Molecule type: mRNA				
A:Residues: 1-266 <MAS>				
A:Cross-references: UNIPROT:Q9ESS3; DDBJ:AB036745				
C:Genetics:				
A:Gene: trb5				
C:Superfamily: X box-binding protein 1; fos/jun DNA-binding domain homology				
C:Keywords: leucine zipper; transcription factor				
Query Match	85.0%;	Score 34;	DB 2;	Length 266;
Best Local Similarity	88.9%;	Pred. No. 32;		
Matches	8;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	LLPLQIILL 9		
DB	187	LLPLQILSL 195		
RESULT 6				
JC4857				
hepatocarcinogenesis-related transcription factor - rat				
C:Species: Rattus norvegicus (Norway rat)				
C:Date: 15-Aug-1996	#sequence_revision 18-Oct-1996	#text_change 09-Jul-2004		
C:Accession: JC4857				
R:Kishimoto, T.; Kokura, K.; Kumagai, Y.; Makino, Y.; Tamura, T. Biochimot. Biophys. Res. Commun. 224, 746-751, 1996				
A:Title: HFG: A b-zip transcription factor that is closely related to the human XBP/TREB				
A:Reference number: JC4857				
A:Accession: JC4857				
A:Molecule type: mRNA				
A:Residues: 1-267 <KIS>				
A:Cross-references: UNIPROT:Q9R1S4				
C:Comment: This is a basic-leucine zipper type transcription factor involved in hepatocel				
C:Genetics:				
A:Gene: htf				
C:Superfamily: X box-binding protein 1; fos/jun DNA-binding domain homology				
C:Keywords: leucine zipper; transcription factor				
F:58-98/Domain: fos/jun DNA-binding domain homology <FUD>				
F:89-126/Region: leucine zipper motif				
Query Match	85.0%;	Score 34;	DB 2;	Length 267;
Best Local Similarity	88.9%;	Pred. No. 32;		
Matches	8;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	LLPLQIILL 9		
DB	188	LLPLQILSL 196		
RESULT 7				
S33603				
surfactant protein D - bovine				
C:Species: Bos primigenius taurus (cattle)				
C:Date: 02-Dec-1993	#sequence_revision 01-Sep-1995	#text_change 17-Mar-1999		
C:Accession: S33603				
R:Lim, B.L.; Lu, J.; Reid, K.B.M. Immunology 78, 159-165, 1993				
A:Title: Structural similarity between bovine conglutinin and bovine lung surfactant prot				
A:Reference number: S33603;	MUID:93170856;	PMID:8436402		
A:Accession: S33603				
A>Status: preliminary				
A:Molecule type: mRNA				
A:Residues: 1-369 <LIM>				
C:Superfamily: pulmonary surfactant protein D; C-type lectin homology				
F:248-367/Domain: C-type lectin homology <LCH>				
Query Match	85.0%;	Score 34;	DB 2;	Length 369;

Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLPLQILL 9  
|||:|  
Db 3 LLPLSVLL 11

RESULT 8  
JN0450  
conglutinin precursor - bovine  
N:Alternate names: C3b-binding protein  
N:Contains: conglutinin-N  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
R:Accession: JN0450; JG2396; S33235; A23740; S36879; S35044; I46010; A29416; S34054  
R:Suzuki, Y.; Yin, Y.P.; Makino, M.; Kurimura, T.; Wakamiya, N.  
Biochem. Biophys. Res. Commun. 191, 335-342, 1993  
A:Title: Cloning and sequencing of a cDNA coding for bovine conglutinin.  
A:Reference number: JN0450; MUID:9313261; PMID:8460993  
A:Accession: JN0450  
A:Molecule type: mRNA  
A:Residues: 1-371 <SUZ>  
A:Cross-references: UNIPROT:P23805; DDBJ:D14085; NID:G285643; PIDN:BAA03170.1; PID:G2856  
A:Experimental source: liver  
R:Kawasaki, N.; Itoh, N.; Kawasaki, T.  
Biochem. Biophys. Res. Commun. 198, 597-604, 1994  
A:Title: Gene organization and 5'-flanking region sequence of conglutinin: A C-type mamm  
A:Reference number: JG2396; MUID:94128104; PMID:8297370  
A:Accession: JG2396  
A:Molecule type: mRNA  
A:Residues: 1-371 <KA2>  
A>Note: The authors translated the codon GAT for residues 250 and 270 as Glu  
R:Lu, J.; Laursen, S.B.; Thiel, S.; Jensenius, J.C.; Reid, K.B.M.  
Biochem. J. 292, 157-162, 1993  
A:Title: The cDNA cloning of conglutinin and identification of liver as a primary site o  
A:Reference number: S33235; MUID:93277452; PMID:7684896  
A:Accession: S33235  
A:Molecule type: mRNA  
A:Residues: 1-172, 'H', 174-217, 'A', 219-271, 'V', 273-371 <LUJ>  
A:Cross-references: EMBL:X71774; NID:G395267; PIDN:CAA50665.1; PID:G395268  
A:Experimental source: liver  
R:Lee, Y.M.; Leiby, K.R.; Allar, J.; Paris, K.; Lerch, B.; Okarna, T.B.  
J. Biol. Chem. 266, 2715-2723, 1991  
A:Title: Primary structure of bovine conglutinin, a member of the C-type animal lectin f  
A:Reference number: A23740; MUID:91131556; PMID:1993651  
A:Accession: A23740  
A:Molecule type: protein  
A:Residues: 21-209, 'S', 211-371 <LEE>  
R:Kawasaki, N.; Yokota, Y.; Kawasaki, T.  
Arch. Biochem. Biophys. 305, 533-540, 1993  
A:Title: Differentiation of conglutination activity and sugar-binding activity of conglu  
A:Reference number: S36879; MUID:93384312; PMID:8373191  
A:Accession: S36879  
A:Molecule type: protein  
A:Residues: 21-54/75-86, 'X', 88-89, 'X', 91, 'X', 93-94/208-209, 'X', 211-227 <KAW>  
A:Experimental source: serum  
R:Lu, J.; Wiedemann, H.; Holmskov, U.; Thiel, S.; Timpl, R.; Reid, K.B.M.  
Eur. J. Biochem. 215, 793-799, 1993  
A:Title: Structural similarity between lung surfactant protein D and conglutinin. Two di  
A:Reference number: S35044; MUID:93358905; PMID:8354286  
A:Accession: S35044  
A:Molecule type: protein  
A:Residues: 75-86, 'X', 88-89, 'X', 91, 'I' <LUA>  
A:Experimental source: lung  
R:Young, N.M.; Leon, M.A.  
Biochem. Biophys. Res. Commun. 143, 645-651, 1987  
A:Title: The carbohydrate specificity of conglutinin and its homology to proteins in the  
A:Reference number: A29416; MUID:87184551; PMID:3566740  
A:Contents: annotation  
R:Malhotra, R.; Laursen, S.B.; Willis, A.C.; Sim, R.B.  
Biochem. J. 293, 15-19, 1993  
A:Title: Research Communication. Localization of the receptor-binding site in the colled

A:Reference number: S34054; MUID:93319501; PMID:8328957  
A:Contents: annotation  
R:Liou, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarna, T.B.; Tauber, A.I.; Sastry,  
J. Immunol. 153, 173-180, 1994  
A:Title: Bovine conglutinin gene exon structure reveals its evolutionary relationship to  
A:Reference number: I46010; MUID:94267222; PMID:8207234  
A:Accession: I46010  
A>Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-371 <LIO>  
A:Cross-references: EMBL:U06860; NID:G507183; PIDN:AAB60624.1; PID:G514256  
C:Comment: This protein mediates the agglutination of erythrocytes with antibody and com  
C:Comment: This protein is a Ca2+-dependent serum lectin specific for N-acetylglucosamine  
C:Genetics:  
A:Gene: CGN1  
A:Introns: 67/1; 106/1; 142/1; 178/1; 217/1; 245/1  
A:Superfamily: pulmonary surfactant protein D; C-type lectin homology  
C:Keywords: calcium binding; glycoprotein; hydroxylysine; hydroxyproline  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-371/Product: conglutinin #status predicted <MAT>  
F:46-214/Region: collagen-like  
F:75-371/Product: conglutinin-N #status predicted <MA2>  
F:248-369/Domain: C-type lectin homology <LCH>  
F:63, 87, 99, 135, 141, 159, 162, 198, 210/Binding site: carbohydrate (Lys) (covalent) #status p  
F:63, 87, 99, 135, 141, 159, 162, 198, 210/Modified site: 5-hydroxylysine (Lys) #status experim  
F:78, 96, 108, 111, 129, 132, 147, 153, 171, 195/Modified site: 4-hydroxyproline (Pro) #status ex  
F:337/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 85.0%; Score 34; DB 1; Length 371;  
Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLPLQILL 9  
|||:|  
Db 3 LLPLSVLL 11

RESULT 9  
I45878  
conglutinin - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: I45878  
R:Liou, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarna, T.B.; Tauber, A.I.; Sastry,  
Gene 141, 277-281, 1994  
A:Title: Bovine conglutinin (BC) mRNA expressed in liver: cloning and characterization of  
A:Reference number: I45878; MUID:94215917; PMID:8163202  
A:Accession: I45878  
A>Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: mRNA  
A:Residues: 1-371 <LIO>  
A:Cross-references: UNIPROT:P23805; GB:L18871; NID:G495012; PIDN:AAA20126.1; PID:G495013  
C:Superfamily: pulmonary surfactant protein D; C-type lectin homology  
F:248-369/Domain: C-type lectin homology <LCH>

Query Match 85.0%; Score 34; DB 2; Length 371;  
Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLPLQILL 9  
|||:|  
Db 3 LLPLSVLL 11

RESULT 10  
TI7215  
hypothetical protein DKFZp434H2235.1 - human  
C:Species: Homo sapiens (man)  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: TI7215  
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A:Reference number: Z18723

A;Accession: T17215  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-429 <POU>  
A;Cross-references: EMBL:AL117404  
A;Experimental source: adult testis; clone DKFZp434H2235  
C;Note: DKFZp434H2235.1

Query Match 85.0%; Score 34; DB 2; Length 429;  
Best Local Similarity 87.5%; Pred. No. 51;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLPLQILL 9  
|||:|  
Db 381 LPLQLLLL 388

RESULT 11  
H82994  
probable permealase of ABC iron transporter PAS216 [imported] - Pseudomonas aeruginosa (strain ATCC 27802)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: H82994  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A;Reference number: AB2950; MUID:20437337; PMID:10984043  
A;Accession: H82994  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-539 <STO>  
A;Cross-references: UNIPROT:Q9HTX4; GB:AE004934; GB:AE004091; NID:99951515; PIDN:AAG0860  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PAS216  
C;Superfamily: sfuB protein

Query Match 85.0%; Score 34; DB 2; Length 539;  
Best Local Similarity 77.8%; Pred. No. 63;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
|||:|  
Db 21 LLPLSVLLL 29

RESULT 12  
T12520  
hypothetical protein DKFZp434G173.1 - human  
C;Species: Homo sapiens (man)  
C;Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 23-Jul-1999  
C;Accession: T12520  
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z17524  
A;Accession: T12520  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1092 <WAM>  
A;Cross-references: EMBL:AL080133  
A;Experimental source: adult testis; clone DKFZp434G173  
C;Genetics:  
A;Note: DKFZp434G173.1

Query Match 85.0%; Score 34; DB 2; Length 1092;  
Best Local Similarity 87.5%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLPLQILL 9  
|||:|

Db 1044 LPLQLLLL 1051

## RESULT 13

T50904  
Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus  
C;Species: Rubrivivax gelatinosus  
C;Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 09-Jul-2004  
C;Accession: T50904  
R;Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.  
submitted to the EMBL Data Library, November 1999  
A;Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynthet  
A;Reference number: Z25270  
A;Accession: T50904  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1236 <NAG>  
A;Cross-references: UNIPROT:Q9JPA4; EMBL:AB034704; PIDN:BAA94057.1  
A;Experimental source: strain IL144  
C;Genetics:  
A;Gene: bchH  
C;Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 85.0%; Score 34; DB 2; Length 1236;  
Best Local Similarity 88.9%; Pred. No. 1.4e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
|||:|  
Db 946 LLPLQIKLL 954

RESULT 14  
H75201  
hypothetical protein PAB0088 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: H75201  
R;anonymous; Genoscope  
submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc  
A;Reference number: A75001  
A;Accession: H75201  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-146 <KAW>  
A;Cross-references: UNIPROT:Q9V2D5; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CAB4906;  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB0088  
C;Superfamily: Pyrococcus abyssi hypothetical protein PAB0088

Query Match 82.5%; Score 33; DB 2; Length 146;  
Best Local Similarity 66.7%; Pred. No. 27;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
|||:|  
Db 68 LLPLQIIIV 76

RESULT 15  
H84314  
cytochrome a3 controlling protein [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: H84314  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;  
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: H84314

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-277 <STO>

A:Cross-references: UNIPROT:Q9HPI3; GB:AE004437; NID:g10581096; PIDN:AAG19884.1; GSPDB:G

C:Genetics:

A:Gene: CCP

Query Match 82.5%; Score 33; DB 2; Length 277;  
Best Local Similarity 75.0%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 8

Db |||||:|

98 LLPLQVIL 105

Search completed: July 13, 2005, 17:31:30

Job time : 14.5556 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-35

Perfect score: 40

Sequence: 1 LLPLQILL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	40	100.0	66	Q6DTY1 homo sapien
2	40	100.0	253	P49862 homo sapien
3	40	100.0	253	Q8N5N9
4	38	95.0	73	Q8VCA9
5	38	95.0	104	Q8R5D6
6	38	95.0	138	Q6PKE2
7	38	95.0	201	FK11 MOUSE
8	38	95.0	361	Q77755 trichosurus
9	37	92.5	552	Q75E55 ashbya goes
10	36	90.0	79	Q8VB59
11	36	90.0	637	Q7RQ54
12	36	90.0	845	Q97H76
13	35	87.5	181	Q6NB96
14	35	87.5	183	Q8MZZ0
15	35	87.5	208	Q21527
16	35	87.5	233	Q8IMF1
17	35	87.5	493	Q92KQ5
18	35	87.5	496	Q25600
19	35	87.5	1620	Q622J8
20	34	85.0	182	Q8MZY6
21	34	85.0	182	Q8MZX7
22	34	85.0	182	Q8MZY6
23	34	85.0	182	Q8MZY6
24	34	85.0	205	Q653Z3
25	34	85.0	242	Q7VHT6
26	34	85.0	256	Q222G5
27	34	85.0	262	Q8BMM1
28	34	85.0	266	Q9ESS3
29	34	85.0	267	Q9R1S4
30	34	85.0	267	Q35426
31	34	85.0	283	Q8K147

32	34	85.0	331	2	Q7X151	Q7X151 oryza sativ
33	34	85.0	338	2	Q6US81	Q6US81 spodoptera
34	34	85.0	369	1	P8FD_BOVIN	P3246 bos taurus
35	34	85.0	369	2	Q863A1	Q863A1 bos taurus
36	34	85.0	371	1	CL46_BOVIN	Q8H29 bos taurus
37	34	85.0	371	1	CONG_BOVIN	P23805 bos taurus
38	34	85.0	489	2	Q8N7G8	Q8N7G8 homo sapien
39	34	85.0	538	2	Q88A51	Q88A51 pseudomonas
40	34	85.0	539	2	Q8HTX4	Q8HTX4 pseudomonas
41	34	85.0	540	2	Q88C16	Q88C16 pseudomonas
42	34	85.0	559	2	Q82114	Q82114 streptomyces
43	34	85.0	668	2	Q86YP9	Q86YP9 homo sapien
44	34	85.0	696	2	Q7VB38	Q7VB38 prochloroco
45	34	85.0	899	2	Q6MZP0	Q6MZP0 homo sapien

#### ALIGNMENTS

##### RESULT 1

Q6DTY1 PRELIMINARY; PRT; 66 AA.  
AC Q6DTY1  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Kallikrein 7 splice variant 3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Kishi T., Michael I.P., Diamandis E.P.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY646152; AAT65047.1;  
SQ SEQUENCE 66 AA; 7171 MW; 82E1C392BC8222FDB CRC64;

Query Match 100.0%; Score 40; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 6 LLPLQILL 14

##### RESULT 2

KLK7\_HUMAN STANDARD; PRT; 253 AA.  
ID KLK7\_HUMAN  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).  
GN Name=KLK7; Synonyms=PRSS6, SCCE;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
RC TISSUE=Skin;  
RX MEDLINE=94308225; PubMed=8034709;  
RA Hansson L., Stromqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;  
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";  
RL J. Biol. Chem. 269:19420-19426(1994).  
RN [2]  
RP SEQUENCE FROM N.A.

RC TISSUE=Keratinocytes;  
 RA Yousef G.M., Scorilas A., Diamandis E.P.;  
 RT "Molecular characterization, mapping and tissue expression of the  
 RT human stratum corneum chymotryptic enzyme gene";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RX PubMed:11054574; DOI=10.1016/S0378-1119(00)00382-6;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepers B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130(2000).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
 RA Wallbrandt P., Egelrud T.;  
 RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
 RT mice; a model for chronic itchy dermatitis";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN CHARACTERIZATION.  
 RX MEDLINE=95314630; PubMed=7794273;  
 RA Skjott A., Stroemqvist M., Egelrud T.;  
 RT "Primary substrate specificity of recombinant human stratum corneum  
 RT chymotryptic enzyme";  
 RL Biochem. Biophys. Res. Commun. 211:586-589(1995).  
 CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive  
 CC structures in the cornified layer of the skin in the continuous  
 CC shedding of cells from the skin surface. Specific for amino acid  
 CC residues with aromatic side chains in the P1 position. SCCE  
 CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
 CC Phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
 CC activation of precursors to inflammatory cytokines.  
 CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
 CC expressed by keratinocytes in the epidermis. Very low levels are  
 CC also seen in the brain and kidney.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
 CC subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L33404; AAC37551.1; -.  
 DR EMBL; AF166330; AAD49718.1; -.  
 DR EMBL; AF243527; AAG33360.1; -.  
 DR EMBL; AF325593; AAK59624.1; -.  
 DR PIR; A53968; A53968.  
 DR HSSP; P00760; 1EZX.  
 DR MEROPS; S01.300; -.  
 DR Genew; HGNC:6368; KLK7.  
 DR H-InvDB; HIX0015373; -.  
 DR MIM; 604438; -.  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR GO; GO:0008544; P:epidermal differentiation; TAS.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
 FT SIGNAL 1 22  
 FT PROPEP 23 29 Activation peptide.

FT CHAIN 30 253 Kallikrein 7.  
 FT ACT\_SITE 70 70 Charge relay system (By similarity).  
 FT ACT\_SITE 112 112 Charge relay system (By similarity).  
 FT ACT\_SITE 205 205 Charge relay system (By similarity).  
 FT DISULFID 36 137 By similarity.  
 FT DISULFID 55 71 By similarity.  
 FT DISULFID ? 239 By similarity.  
 FT DISULFID 144 211 By similarity.  
 FT DISULFID 176 190 By similarity.  
 FT DISULFID 201 226 By similarity.  
 FT CARBOHYD 246 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;  
 Query Match 100.0%; Score 40; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLPLQILL 9  
 Db 6 LLPLQILL 14  
 RESULT 3  
 Q8NSN9 ID Q8NSN9 PRELIMINARY; PRT; 253 AA.  
 AC Q8NSN9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DE Stratum corneum chymotryptic enzyme, preproprotein.  
 GN Name=KLK7;  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; BC032005; AAH32005.1; -.  
 DR HSSP; P00760; 1EZX.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.



```

DR InterPro: IPR009003; Pept_Ser_Cys.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQLLLL 9
DB 6 LLPLQLLLL 14

RESULT 4
ID Q8VCA9 PRELIMINARY; PRT; 73 AA.
AC Q8VCA9;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Fkbp11 protein.
GN Name=Fkbp11;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX STRAIN=CZSCH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the FkBP-type tripease family.
DR EMBL; BC022900; AAH22900.1; -.
DR HSP; P18203; 1PKL.
DR MGD; MGI:1913370; Fkbp11.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR001179; FkBP_PPase.
DR Pfam; PF00254; FkBP_C; 1.
DR PROSITE; PS00059; FkBP_PPase; 1.
SQ SEQUENCE 104 AA; 11095 MW; 0534D57467566914 CRC64;

Query Match 95.0%; Score 38; DB 2; Length 104;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQLLLL 9
DB 7 LLPLQLLLL 15

RESULT 6
ID Q6PKB2 PRELIMINARY; PRT; 138 AA.
AC Q6PKB2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Fkbp11 protein.

```

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FBV/N; TISSUE=Mammary tumor; DOI=10.1073/pnas.242603899;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Schaeetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FBV/N; TISSUE=Mammary tumor;  
 RX STRAUSBERG R.;  
 RA EMBL: BC002311; AA02311.1;  
 DR GO: GO:0005615; C:extracellular space; TAS.  
 DR GO: GO:0016021; C:integral to membrane; TAS.  
 DR InterPro: IPR001179; FKBP\_PPIase.  
 DR PROSITE: PS0059; FKBP\_PPIASE; 1.  
 SQ SEQUENCE 138 AA; 15105 MW; C138B8B0EFD959D CRC64;  
 Query Match 95.0%; Score 38; DB 2; Length 138;  
 Best Local Similarity 88.9%; Pred. No. 14;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQILL 9  
 D5 7 LLPLQLLL 15  
 RESULT 7  
 ID FK11 MOUSE STANDARD; PRT; 201 AA.  
 AC Q9DLM7; Q9CRE4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FK506 binding protein 11 precursor (EC 5.2.1.8) (Peptidyl-prolyl cis-  
 DE trans isomerase) (PPIase) (Rotamase) (19 kDa FK506-binding protein)  
 DE (FKBP-19).  
 GN Name=FKBP11;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FBV/N; TISSUE=Embryo;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Otsu N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fietzer C.F., Forrest A., Frazer K.S.,  
 RA Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Schaeetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: PPIases accelerate the folding of proteins during  
 CC protein synthesis.  
 CC -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline  
 CC (omega=0).  
 CC -!- SIMILARITY: Belongs to the FKBP-type PPIase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AK003331; BAB22719.1; -  
 CC EMBL: AK019132; BAB31559.1; -  
 CC EMBL: BC037596; AAH37596.1; -  
 CC HSSP: P20071; ITCO.  
 CC MGP: MG1:1913370; FKBP11.  
 DR InterPro: IPR001179; FKBP\_PPIase.  
 DR Pfam: PF00254; FKBP\_C; 1.  
 DR PROSITE: PS0059; FKBP\_PPIASE; 1.  
 KW Isomerase; Rotamase; Signal. Potential.  
 FT SIGNAL 1 27  
 FT CHAIN 28 201 FK506 binding protein 11.

FT DOMAIN 57 144 PPIase, PKBP-type.  
 FT CONFLICT 53 S -> F (in Ref. 1; BAB31559).  
 FT CONFLICT 198 S -> R (in Ref. 1; BAB31559).  
 SQ SEQUENCE 201 AA; 22137 MW; 949555C57264BD82 CRC64;  
 Query Match 95.0%; Score 38; DB 1; Length 201;  
 Best Local Similarity 88.9%; Pred. No. 20;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQILL 9  
 |||||:  
 Db 7 LLPLQLLL 15  
 |||||:  
 RESULT 8  
 IHA TRIVU STANDARD; PRT; 361 AA.  
 AC Q77755;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Inhibin alpha chain precursor.  
 GN Name=INHA;  
 OS Trichosurus vulpecula (Brush-tailed possum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.  
 OX NCBI\_TaxID=9337;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99027340; PubMed=9801457;  
 RA Vanmontfort D., Fidler A.E., Heath D.A., Lawrence S.B., Tisdall D.J.,  
 RA Greenwood P.J., McNatty K.;  
 RT "cDNA sequence analysis, gene expression and protein localisation of  
 the inhibin alpha subunit of Australian brushtail possum (Trichosurus  
 vulpecula).";  
 RL J. Mol. Endocrinol. 21:141-152(1998).  
 CC -!- FUNCTION: Inhibins and activins inhibit and activate.  
 CC respectively, the secretion of follitropin by the pituitary gland.  
 CC Inhibins/activins are involved in regulating a number of diverse  
 CC functions such as hypothalamic and pituitary hormone secretion,  
 CC gonadal hormone secretion, germ cell development and maturation,  
 CC erythroid differentiation, insulin secretion, nerve cell survival,  
 CC embryonic axial development or bone growth, depending on their  
 CC subunit composition. Inhibins appear to oppose the functions of  
 CC activins.  
 CC -!- SUBUNIT: Dimeric, linked by one or more disulfide bonds. Inhibin A  
 CC is a dimer of alpha and beta-A. Inhibin B is a dimer of alpha and  
 CC beta-B.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF033340; AAC63945.1;  
 CC GO; GO:0005576; C:extracellular; ISS.  
 CC GO; GO:0017105; F:activin inhibitor activity; ISS.  
 CC GO; GO:0005125; F:cytokine activity; ISS.  
 CC GO; GO:0008083; F:growth factor activity; ISS.  
 CC GO; GO:0005179; F:hormone activity; ISS.  
 CC GO; GO:0005515; F:protein binding; ISS.  
 CC GO; GO:0007050; P:cell cycle arrest; ISS.  
 CC GO; GO:0030154; P:cell growth and/or maintenance; ISS.  
 CC GO; GO:0008151; P:cell growth and/or maintenance; ISS.  
 CC GO; GO:0007166; P:cell surface receptor linked signal transdu. .; ISS.  
 CC GO; GO:0007267; P:cell-cell signaling; ISS.  
 CC GO; GO:0030218; P:erythrocyte differentiation; ISS.  
 CC GO; GO:0042541; P:hemoglobin biosynthesis; ISS.  
 CC GO; GO:0006917; P:induction of apoptosis; ISS.

DR GO; GO:0045578; P:negative regulation of B-cell differentiation; ISS.  
 DR GO; GO:0045786; P:negative regulation of cell cycle; ISS.  
 DR GO; GO:0046882; P:negative regulation of follicle-stimulating. .; ISS.  
 DR GO; GO:0045077; P:negative regulation of interferon-gamma bio. .; ISS.  
 DR GO; GO:0045650; P:negative regulation of macrophage different. .; ISS.  
 DR GO; GO:0042326; P:negative regulation of phosphorylation; ISS.  
 DR GO; GO:0007399; P:neurogenesis; ISS.  
 DR GO; GO:0001541; P:ovarian follicle development; ISS.  
 DR GO; GO:0046881; P:positive regulation of follicle-stimulating. .; ISS.  
 DR InterPro; IPR003405; Inhibin\_alpha.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF\_beta; 1.  
 DR PRINTS; PR00669; INHIBIN.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Glycoprotein; Growth factor; Hormone; Signal.  
 FT SIGNAL 1 21  
 FT PROPEP 22 230  
 FT CHAIN 231 361  
 FT DISULFID 260 323  
 FT DISULFID 289 358  
 FT DISULFID 293 360  
 FT DISULFID 322 322  
 FT CARBOHYD 48 48  
 FT CARBOHYD 144 144  
 FT CARBOHYD 266 266  
 SQ SEQUENCE 361 AA; 38945 MW; D661CDF93CDA87D CRC64;  
 Query Match 95.0%; Score 38; DB 1; Length 361;  
 Best Local Similarity 88.9%; Pred. No. 36;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQILL 9  
 |||||:  
 Db 5 LLPLQLLL 13  
 |||||:  
 RESULT 9  
 QY75E55 PRELIMINARY; PRT; 552 AA.  
 AC Q75E55;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE ABL185Cp.  
 GN ORFNames=ABL185C;  
 OS Ashbya gossypii (Yeast) (Bremothecium gossypii).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Bremothecium.  
 OX NCBI\_TaxID=33169;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 10895;  
 RA Voegel S.E., Dietrich F.S., Brachat S., Lerch A., Gaffney T.,  
 RA Philippsen P.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE016815; AAS50586.1;  
 DR AGD; ABL185C;  
 DR InterPro; IPR005599; Alg9 trans.  
 DR InterPro; IPR001202; WW\_Rp5\_WWP.  
 DR Pfam; PF03901; Glyco\_transf\_22; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; UNKNOWN 1.  
 SQ SEQUENCE 552 AA; 62707 MW; 62A8842023DEEA04 CRC64;  
 Query Match 92.5%; Score 37; DB 2; Length 552;  
 Best Local Similarity 77.8%; Pred. No. 87;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLPLQILL 9  
 |||||:  
 Db 67 LMPLOVLL 75  
 |||||:

```
RESULT 10
Q8VB59 PRELIMINARY; PRT; 79 AA.
AC Q8VB59;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Wsv132 (WSSV187).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae; Whispovirus.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21548311; PubMed=11689662;
RX DOI=10.1128/JVI.75.23.11811-11820.2001;
RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
RT virus.";
RL J. Virol. 75:11811-11820(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517548; PubMed=11062040; DOI=10.1006/viro.2000.0597;
RX Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RX Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT "Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT cellular-type thymidine kinase and thymidylate kinase.";
RL Virology 277:100-110(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21844071; PubMed=11853398; DOI=10.1006/viro.2001.1273;
RX Chen L.F., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RX Lo C.F., Kou G.H.;
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT spot syndrome virus and characterization of the motif important for
RT targeting VP35 to the nuclei of transfected insect cells.";
RL Virology 293:44-53(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX Lo C.-F., Kou G.-H.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332093; AAL33136.1; -.
DR EMBL; AF440570; AAL89055.1; -.
SQ SEQUENCE 79 AA; 9287 MW; 800910C7FFA0EC7E CRC64;

Query Match 90.0%; Score 36; DB 2; Length 79;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPLQLILL 9
Db 34 LPLQLILL 42

RESULT 11
Q7RQ54 PRELIMINARY; PRT; 637 AA.
AC Q7RQ54;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mtn3/saliva family, putative.
GN Name=PY01249;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RX Carlton J.M., Angiolini S.V., Suh B.B., Kooij T.W., Perteu M.,
RX Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,

Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
Carucci D.J.;
"Genome sequence and comparative analysis of the model rodent malaria
parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000328; EAA20540.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR004316; MCN3_slv.
DR Pfam; PF03083; MCN3_slv; 2.
SQ SEQUENCE 637 AA; 70992 MW; BC506320B2AB1FB5 CRC64;

Query Match 90.0%; Score 16; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPLQLILL 9
Db 7 LPLQLILL 14

RESULT 12
Q97H76 PRELIMINARY; PRT; 845 AA.
AC Q97H76;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cation transport P-type ATPase.
GN OrderedLocusNames=CAC2137;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RX Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RX Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RX Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RX Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007714; AAK80095.1; -.
DR PIR; D97163; D97163.
DR HSSP; P04191; 1SU4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005562; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. .; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0015932; P:proton transport; IEA.
DR InterPro; IPR001757; ATPase_EI-E2.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehal_Like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
```

```

DR Pfam: PF00122; E1-E2 ATPase; 1.
DR Pfam: PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATALPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 845 AA; 93779 MW; 401293AA1FF9D757 CRC64;

Query Match      90.0%; Score 36; DB 2; Length 845;
Best Local Similarity 77.8%; Pred. No. 2.1e-02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9
DB 670 LLPIQLLI 678

RESULT 13
Q6NB96 PRELIMINARY; PRT; 181 AA.
AC Q6NB96;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=RPA0932;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J.E., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL: BX52595; CAE26376.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 181 AA; 19632 MW; CDB4D8210E565D1 CRC64;

Query Match      87.5%; Score 35; DB 2; Length 181;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 8
DB 24 LLPLQVLL 31

RESULT 14
Q8MZZO PRELIMINARY; PRT; 183 AA.
AC Q8MZZO;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Acyl-CoA desaturase Hasappae (Fragment).
OS Helicoverpa assulta (Oriental tobacco budworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Helicoverpa.
OX NCBI_TaxID=52344;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22412134; PubMed=12524345;
RA Knipple D.C., Rosenfield C.L., Nielsen R., You K.M., Jeong S.E.;
RT "Evolution of the integral membrane desaturase gene family in moths

and flies.";
RL Genetics 162:1737-1752(2002).
DR EMBL; AF482910; AAM28485.1; -.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004768; F:stearyl-CoA 9-desaturase activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001522; Desaturase.
DR InterPro; IPR005804; Fa_desat.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00487; FA_desaturase; 1.
DR PRINTS; PR00075; FACDSDATRASE.
DR PRINTS; PR00024; HOMEBOX.
FT NON_TER 1
FT NON_TER 183
SQ SEQUENCE 183 AA; 21196 MW; 728DA94FAFCD14FA CRC64;

Query Match      87.5%; Score 35; DB 2; Length 183;
Best Local Similarity 87.5%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLPLQILL 9
DB 14 LLPLQVLL 21

RESULT 15
Q21527 PRELIMINARY; PRT; 208 AA.
AC Q21527;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 4 (fragment).
GN Name=ND4;
OS Clethrionomys gapperi (Southern red-backed vole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
OC Clethrionomys.
OX NCBI_TaxID=56223;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9815303; PubMed=9491603;
RA Engel S.R., Hogan K.M., Taylor J.F., Davis S.K.;
RT "Molecular systematics and paleobiogeography of the South American
RT sigmodontine rodents.";
RL Mol. Biol. Evol. 15:35-49(1998).
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; U83808; AAB87168.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR003918; NADHub_oxred4.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR000260; Oxidored_q5_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR PRINTS; PR01437; NUOXDRDTASE4.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
FT NON_TER 208
SQ SEQUENCE 208 AA; 23967 MW; 8AF1788697AED6A2 CRC64;

Query Match      87.5%; Score 35; DB 2; Length 208;
Best Local Similarity 77.8%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

Qv 1 LPLQILL 9  
|:|||||:  
Db 99 LVPLQILLI 107

Search completed: July 13, 2005, 17:29:24  
Job time : 64.2222 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-35

Perfect score: 40

Sequence: 1 LLPLOIILL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1 pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	9	3 US-09-502-600-35	Sequence 35, Appl
2	40	100.0	9	4 US-09-918-243-35	Sequence 35, Appl
3	40	100.0	253	2 US-08-557-146-2	Sequence 2, Appli
4	40	100.0	253	2 US-08-824-874-3	Sequence 3, Appli
5	40	100.0	253	2 US-09-154-344-2	Sequence 2, Appli
6	40	100.0	253	3 US-08-930-188-2	Sequence 2, Appli
7	40	100.0	253	3 US-09-210-084-3	Sequence 3, Appli
8	40	100.0	253	4 US-09-764-762-3	Sequence 3, Appli
9	40	100.0	253	5 PCT-US96-04294-2	Sequence 2, Appli
10	40	100.0	265	4 US-09-949-016-7716	Sequence 7716, Ap
11	36	90.0	9	3 US-09-502-600-33	Sequence 33, Appl
12	36	90.0	9	4 US-09-918-243-33	Sequence 33, Appl
13	34	85.0	73	4 US-09-205-258-1083	Sequence 1083, Ap
14	34	85.0	302	4 US-09-393-634-37	Sequence 37, Appl
15	34	85.0	424	4 US-09-205-258-1079	Sequence 1079, Ap
16	34	85.0	569	4 US-09-252-991A-22870	Sequence 22870, A
17	34	85.0	812	4 US-09-489-039A-12075	Sequence 12075, A
18	33	82.5	180	4 US-09-270-767-42872	Sequence 42872, A
19	33	82.5	190	1 US-08-339-152A-19	Sequence 19, Appl
20	33	82.5	190	2 US-08-007-999B-6	Sequence 6, Appli
21	33	82.5	190	2 US-08-689-276A-6	Sequence 6, Appli
22	33	82.5	201	4 US-09-919-039-311	Sequence 311, App
23	33	82.5	232	2 US-08-956-047-36	Sequence 36, Appl
24	33	82.5	634	1 US-08-339-152A-17	Sequence 17, Appl
25	33	82.5	653	1 US-08-339-152A-16	Sequence 16, Appl
26	33	82.5	653	2 US-08-007-999B-3	Sequence 3, Appli
27	33	82.5	653	2 US-08-689-276A-3	Sequence 3, Appli

28 82.5 2787 3 US-09-245-041-15 Sequence 15, Appl  
29 82.5 2787 4 US-09-358-055B-15 Sequence 15, Appl  
30 82.5 2787 4 US-09-893-238-15 Sequence 15, Appl  
31 80.0 9 3 US-09-502-600-36 Sequence 36, Appl  
32 80.0 9 4 US-09-918-243-36 Sequence 36, Appl  
33 80.0 12 3 US-09-502-600-28 Sequence 28, Appl  
34 80.0 12 4 US-09-918-243-28 Sequence 28, Appl  
35 80.0 135 4 US-09-270-767-37619 Sequence 37619, A  
36 80.0 135 4 US-09-270-767-52836 Sequence 52836, A  
37 80.0 142 4 US-09-248-796A-21399 Sequence 21399, A  
38 80.0 264 3 US-08-713-556F-42 Sequence 42, Appl  
39 80.0 269 3 US-08-713-556F-38 Sequence 38, Appl  
40 80.0 313 3 US-08-713-556F-36 Sequence 36, Appl  
41 80.0 395 4 US-09-489-039A-9507 Sequence 9507, Ap  
42 80.0 402 2 US-08-477-254A-2 Sequence 2, Appli  
43 80.0 402 2 US-08-472-576B-2 Sequence 2, Appli  
44 80.0 402 2 US-08-428-734B-2 Sequence 2, Appli  
45 80.0 402 3 US-09-063-237-1 Sequence 1, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-502-600-35

; Sequence 35, Application US/09502600A

; Patent No. 6294344

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of Ovarian Cancer

; FILE REFERENCE: D6223CIP-C

; CURRENT FILING DATE: 2000-02-11

; CURRENT APPLICATION NUMBER: US/09/502,600A

; PRIOR APPLICATION NUMBER: 09/039,211

; PRIOR FILING DATE: 03-14-1998

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 35

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Residues 6-14 of the SCCE protein

US-09-502-600-35

Query Match 100.0%; Score 40; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLOIILL 9

DB 1 LLPLOIILL 9

##### RESULT 2

US-09-918-243-35

; Sequence 35, Application US/09918243

; Patent No. 6627403

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; APPLICANT: Cannon, Martin J.

; APPLICANT: Santin, Alessandro

; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP

; CURRENT APPLICATION NUMBER: US/09/918,243

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US

; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 35

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-918-243-35

Query Match      100.0%; Score 40; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9
    |||||
Db 1 LLPLQILL 9

RESULT 3
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hanson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2

Query Match      100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9
    |||||
Db 6 LLPLQILL 14

RESULT 4
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-08-824-874-3

Query Match      100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9
    |||||
Db 6 LLPLQILL 14

RESULT 5
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hanson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```



APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-154-344-2

Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9  
DB 6 LLPLQILL 14

RESULT 6  
US-08-930-188-2  
Sequence 2, Application US/08930188  
Patent No. 6093397  
GENERAL INFORMATION:  
APPLICANT: Dixon, Eric P.  
ADDRESSEE: Johnstone, Edward M.  
APPLICANT: Little, Sheila P.  
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
RELATED NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930,188  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/416,257  
FILING DATE: 04-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Blalock, Donna K.  
REGISTRATION NUMBER: 38,082  
REFERENCE/DOCKET NUMBER: X9239  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-277-1090  
TELEFAX: 317-276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-930-188-2

Query Match 100.0%; Score 40; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLPLQILL 9  
DB 6 LLPLQILL 14  
RESULT 7  
US-09-210-084-3  
Sequence 3, Application US/09210084  
Patent No. 6197511  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,084  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,874  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0252 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 532504  
US-09-210-084-3

Query Match 100.0%; Score 40; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 9  
DB 6 LLPLQILL 14

RESULT 8  
US-09-764-762-3  
Sequence 3, Application US/09764762  
Patent No. 6472195  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5

```
/
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/764,762
/ FILING DATE: 16-Jan-2001
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/210,084
/ FILING DATE: <Unknown>
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0252 US
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/
/ INFORMATION FOR SEQ ID NO: 3:
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 253 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 532504
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

      Query Match      100.0%; Score 40; DB 4; Length 253;
      Best Local Similarity 100.0%; Pred. No. 4.2;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9
Db 6 LLPLQILL 14

RESULT 9
PCT-US96-04294-2
/ Sequence 2, Application PC/TUS9604294
/ GENERAL INFORMATION:
/ APPLICANT: Dixon, Eric P.
/ APPLICANT: Johnstone, Edward M.
/ APPLICANT: Little, Sheila P.
/ TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
/ TITLE OF INVENTION: RELATED NUCLEIC ACIDS
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eli Lilly and Company
/ STREET: Lilly Corporate Center
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: United States of America
/ ZIP: 46285
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/04294
/ FILING DATE:
/ CLASSIFICATION:
```

```
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/416,257
/ FILING DATE: 04-APR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Blalock, Donna K.
/ REGISTRATION NUMBER: 38,082
/ REFERENCE/DOCKET NUMBER: X9239
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-277-1090
/ TELEFAX: 317-276-3861
/
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 253 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ PCT-US96-04294-2

      Query Match      100.0%; Score 40; DB 5; Length 253;
      Best Local Similarity 100.0%; Pred. No. 4.2;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9
Db 6 LLPLQILL 14

RESULT 10
US-09-949-016-7716
/ Sequence 7716, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7716
/ LENGTH: 265
/ TYPE: PRT
/ ORGANISM: Human
/
/ US-09-949-016-7716

      Query Match      100.0%; Score 40; DB 4; Length 255;
      Best Local Similarity 100.0%; Pred. No. 4.4;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9
Db 18 LLPLQILL 26

RESULT 11
US-09-502-600-33
/ Sequence 33, Application US/09502600A
/ Patent No. 6294344
/ GENERAL INFORMATION:
/ APPLICANT: O'Brien, Timothy J.
/ TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
/ TITLE OF INVENTION: Ovarian Cancer
/ FILE REFERENCE: D6223CIP-C
/ CURRENT FILING DATE: 2000-02-11
/ CURRENT APPLICATION NUMBER: US/09/502,600A
```

```
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-502-600-33

Query Match          90.0%; Score 36; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 8
Db 2 LLPLQILL 9

RESULT 12
US-09-918-243-33
; Sequence 33, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-918-243-33

Query Match          90.0%; Score 36; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQILL 8
Db 2 LLPLQILL 9

RESULT 13
US-09-205-258-1083
; Sequence 1083, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1083
```

```
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-1083

Query Match      85.0%; Score 34; DB 4; Length 73;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPLQLLLL 9
   |||:||||
Db 25 LPLQLLLL 32

RESULT 14
US-09-393-634-37
; Sequence 37, Application US/09393634
; Patent No. 6558910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR02
; NAME/KEY: MOD.RES
; LOCATION: (143)
; OTHER INFORMATION: Xaa = any amino acid
US-09-393-634-37

Query Match      85.0%; Score 34; DB 4; Length 302;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLPLQLLLL 9
   |:||||:
Db 42 LMPIQLLLM 50

RESULT 15
US-09-205-258-1079
; Sequence 1079, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
```

; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1079  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (152)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (314)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (359)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-205-258-1079

Query Match 85.0%; Score 34; DB 4; Length 424;  
Best Local Similarity 87.5%; Pred. No. 97;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPLQILL 9  
|||:|  
Db 376 LPLQLLL 383

Search completed: July 13, 2005, 17:34:27  
Job time : 20.4444 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds

(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-35

Perfect score: 40

Sequence: 1 LLPLQILL 9

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	9 US-09-918-243-35	Sequence 35, Appl
2	40	100.0	9	2 US-09-905-083-35	Sequence 35, Appl
3	40	100.0	9	15 US-10-372-521-35	Sequence 35, Appl
4	40	100.0	9	15 US-10-831-075-35	Sequence 35, Appl
5	40	100.0	198	15 US-10-262-511-96	Sequence 96, Appl
6	40	100.0	250	15 US-10-262-511-92	Sequence 92, Appl
7	40	100.0	253	9 US-09-888-615-98	Sequence 98, Appl
8	40	100.0	253	9 US-09-764-762-3	Sequence 3, Appl
9	40	100.0	253	14 US-10-071-214-2	Sequence 2, Appl
10	40	100.0	253	14 US-10-071-214-48	Sequence 48, Appl
11	40	100.0	253	14 US-10-264-283-90	Sequence 90, Appl

12	40	100.0	253	15	US-10-295-027-498	Sequence 498, Appl
13	40	100.0	253	15	US-10-173-999-48	Sequence 48, Appl
14	40	100.0	253	16	US-10-408-765A-639	Sequence 639, Appl
15	40	100.0	253	16	US-10-643-795A-95	Sequence 95, Appl
16	40	100.0	253	17	US-10-948-518-95	Sequence 95, Appl
17	40	100.0	253	17	US-10-868-490A-1	Sequence 1, Appl
18	40	100.0	257	15	US-10-344-394-38	Sequence 38, Appl
19	38	95.0	229	16	US-10-425-115-227327	Sequence 227327
20	36	90.0	9	9	US-09-918-243-33	Sequence 33, Appl
21	36	90.0	9	9	US-09-905-083-33	Sequence 33, Appl
22	36	90.0	9	15	US-10-372-521-33	Sequence 33, Appl
23	36	90.0	9	16	US-10-831-075-33	Sequence 33, Appl
24	36	90.0	70	15	US-10-424-599-207950	Sequence 207950
25	36	90.0	637	17	US-10-732-923-928	Sequence 928, Appl
26	36	90.0	845	15	US-10-282-122A-51875	Sequence 51875, A
27	36	90.0	868	15	US-10-369-493-10255	Sequence 10255, A
28	35	87.5	23	9	US-09-864-761-46097	Sequence 46097, A
29	35	87.5	58	15	US-10-424-599-234176	Sequence 234176
30	35	87.5	215	16	US-10-425-115-284316	Sequence 284316
31	35	87.5	305	15	US-10-335-977-6098	Sequence 6098, Ap
32	35	87.5	306	15	US-10-335-977-6097	Sequence 6097, Ap
33	35	87.5	329	15	US-10-335-977-6100	Sequence 6099, Ap
34	35	87.5	477	15	US-10-335-977-6100	Sequence 6100, Ap
35	35	87.5	477	15	US-10-335-977-6101	Sequence 6101, Ap
36	35	87.5	480	9	US-09-895-913A-92	Sequence 92, Appl
37	35	87.5	637	15	US-10-425-114-70206	Sequence 70206, A
38	35	87.5	637	16	US-10-425-115-294490	Sequence 294490
39	35	87.5	1498	16	US-10-437-963-121544	Sequence 121544
40	34	85.0	61	16	US-10-425-115-218429	Sequence 218429
41	34	85.0	70	15	US-10-097-065-220	Sequence 220, App
42	34	85.0	70	14	US-10-372-876-220	Sequence 220, App
43	34	85.0	73	10	US-09-933-767-1083	Sequence 1083, Ap
44	34	85.0	73	14	US-10-004-860-1083	Sequence 1083, Ap
45	34	85.0	73	14	US-10-023-282-1083	Sequence 1083, Ap

#### ALIGNMENTS

RESULT 1  
US-09-918-243-35  
; Sequence 35, Application US/09918243  
; Patent No. US2020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 35  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 6-14 of the SCCE protein  
US-09-918-243-35

Query Match 100.0%; Score 40; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLPLQILL 9  
Db 1 LLPLQILL 9

RESULT 2

```
US-09-905-083-35
; Sequence 35, Application US/0905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-905-083-35

Query Match      100.0%; Score 40; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILLL 9
Db 1 LLPLQILLL 9

RESULT 3
US-10-372-521-35
; Sequence 35, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-10-372-521-35

Query Match      100.0%; Score 40; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILLL 9
Db 1 LLPLQILLL 9

RESULT 4
US-10-831-075-35
; Sequence 35, Application US/10831075
; Publication No. US20040224891A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
```

```
; FILE REFERENCE: D6223CIP/C/D/CIP3
; CURRENT APPLICATION NUMBER: US/10/831,075
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 10/372,521
; PRIOR FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-10-831-075-35

Query Match      100.0%; Score 40; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILLL 9
Db 1 LLPLQILLL 9

RESULT 5
US-10-262-511-96
; Sequence 96, Application US/10262511
; Publication No. US20040039223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
```



; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/373,260  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 96  
 ; LENGTH: 198  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-96

Query Match 100.0%; Score 40; DB 15; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
 Db 6 LLPLQILL 14

## RESULT 6

US-10-262-511-92  
 ; Sequence 92, Application US/10262511  
 ; Publication No. US20040038223A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Ju, Jingfang  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Guo, Xiaojia (Sasha)  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Ort, Tatiana  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Zerhusen, Bryan D.  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Pena, Carol E. A.  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Shinkets, Richard A.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Ages, Michele L.  
 ; APPLICANT: Berghs, Constance  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-462C  
 ; CURRENT APPLICATION NUMBER: US/10/262,511  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/373,815  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,642  
 ; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,038  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/373,260  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 92  
 ; LENGTH: 250  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-92

Query Match 100.0%; Score 40; DB 15; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
 Db 3 LLPLQILL 11

## RESULT 7

US-09-888-615-98  
 ; Sequence 98, Application US/09888615  
 ; Patent No. US20020064856A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PLOWMAN, GREGORY  
 ; APPLICANT: WHYTE, DAVID  
 ; APPLICANT: CAENEPEEL, SEAN  
 ; APPLICANT: CHARYDCZAK, GLEN  
 ; APPLICANT: MANNING, GERARD  
 ; APPLICANT: SUDARSANAM, SUCHA  
 ; TITLE OF INVENTION: NOVEL PROTEASES  
 ; FILE REFERENCE: 038602/1214  
 ; CURRENT APPLICATION NUMBER: US/09/888,615  
 ; CURRENT FILING DATE: 2001-06-26  
 ; PRIOR APPLICATION NUMBER: 60/214,047  
 ; PRIOR FILING DATE: 2000-06-26  
 ; NUMBER OF SEQ ID NOS: 150  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 98  
 ; LENGTH: 253  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-888-615-98

Query Match 100.0%; Score 40; DB 9; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
 Db 6 LLPLQILL 14

## RESULT 8

US-09-764-762-3  
 ; Sequence 3, Application US/09764762  
 ; Patent No. US20020068341A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; Lel, Preeti  
 ; TITLE OF INVENTION: NOVEL KALLIKREIN  
 ; NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;; STREET: 3174 Porter Drive  
;; CITY: Palo Alto  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94304  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/764,762  
;; FILING DATE: 16-Jan-2001  
;; CLASSIFICATION: <Unknown>  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/210,084  
;; FILING DATE: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Billings, Lucy J.  
;; REGISTRATION NUMBER: 36,749  
;; REFERENCE/DOCKET NUMBER: PP-0252 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-855-0555  
;; TELEFAX: 415-845-4166  
;;  
;; INFORMATION FOR SEQ ID NO: 3:  
;;  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 253 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: GenBank  
;; CLONE: 532504  
;;  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-764-762-3

Query Match 100.0%; Score 40; DB 9; Length 253;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 6 LLPLQILL 14

RESULT 9  
US-10-071-214-2  
; Sequence 2, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGELRUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-071-214-2

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 6 LLPLQILL 14  
RESULT 10  
US-10-071-214-48  
; Sequence 48, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGELRUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 48  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Deduced amino acid sequence from the C-terminal part of SCCE from  
US-10-071-214-48

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 6 LLPLQILL 14

RESULT 11  
US-10-264-283-90  
; Sequence 90, Application US/10264283  
; Publication No. US20030144494A1  
; GENERAL INFORMATION:  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.590  
; CURRENT APPLICATION NUMBER: US/10/264,283  
; CURRENT FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 90  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-283-90

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLPLQILL 9  
Db 6 LLPLQILL 14

RESULT 12  
US-10-295-027-498  
; Sequence 498, Application US/10295027

```
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natcha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLPLQILL 9
Db      6 LLPLQILL 14

RESULT 13
US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US2004000563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
```

```
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-48

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLPLQILL 9
Db      6 LLPLQILL 14

RESULT 14
US-10-408-765A-639
; Sequence 639, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 639
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-639

Query Match      100.0%; Score 40; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLPLQILL 9
Db      6 LLPLQILL 14

RESULT 15
US-10-643-795A-95
; Sequence 95, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026R1-US
; CURRENT APPLICATION NUMBER: US/10/643,795A
```

```
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 95
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-643-795A-95
```

```
Query Match      100.0%; Score 40; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLPLQILL 9
        |||||||
Db      6 LLPLQILL 14
```

Search completed: July 13, 2005, 18:23:39  
Job time : 66.8889 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083A-36

Perfect score: 40

Sequence: 1 SLLPLQLL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq 16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	4 AAE08241	Aae08241 Human str
2	40	100.0	9	ADR68797	Adr68797 Human str
3	40	100.0	136	4 ABG23378	Abg23378 Novel hum
4	40	100.0	198	6 ADA05736	Ada05736 Human NOV
5	40	100.0	198	8 ADM62900	Adm62900 Human NOV
6	40	100.0	250	6 ADA05732	Ada05732 Human NOV
7	40	100.0	250	8 ADM62896	Adm62896 Human NOV
8	40	100.0	253	2 AAR67888	Aar67888 Human str
9	40	100.0	253	2 AAW05383	Aaw05383 Human any
10	40	100.0	253	5 ABB84421	Abb84421 Human SCC
11	40	100.0	253	5 ABB84406	Abb84406 Human SCC
12	40	100.0	253	5 AAU82740	Aau82740 Amino aci
13	40	100.0	253	6 ABU07440	Abu07440 Protein d
14	40	100.0	253	6 ABR58471	Abu07471 Protein d
15	40	100.0	253	7 ADR58471	Adr58471 Human str
16	40	100.0	253	7 ADR80484	Adr80484 Ovarian c
17	40	100.0	253	7 ADJ68833	Adj68833 Human hea
18	40	100.0	253	7 ADN3180	Adn3180 Cancer/an
19	40	100.0	253	8 ADL06515	Adl06515 Human tum
20	40	100.0	253	8 ADN04182	Adn04182 Antipsori
21	40	100.0	253	8 ADR72880	Adr72880 Human ova
22	40	100.0	257	3 AAB21326	Aab21326 Human HSC
23	36	90.0	9	4 AAE08238	Aae08238 Human str
24	36	90.0	9	4 ADR68794	Adr68794 Human str
25	35	87.5	868	8 ADS21222	Ads21222 Bacterial

26	34	85.0	61	4	AAO12472	Aao12472 Human pol
27	34	85.0	812	7	ABO65558	AbO65558 Klebsiell
28	33	82.5	394	6	ABU17803	Abu17803 Protein e
29	33	82.5	397	7	ADJ71036	Adj71036 Human hea
30	33	82.5	804	6	ABP70828	Abp70828 Murine C1
31	32	80.0	9	4	AAE08240	Aae08240 Human str
32	32	80.0	9	4	AAE08320	Aae08320 Human str
33	32	80.0	9	8	ADR68877	Adr68877 Human str
34	32	80.0	9	8	ADR68796	Adr68796 Human str
35	32	80.0	71	4	AAW89706	Aaw89706 Human imm
36	32	80.0	73	4	ABG19347	Abg19347 Novel hum
37	32	80.0	78	8	ADL181901	Adl181901 P. aerugi
38	32	80.0	139	7	ADB65713	Adb65713 Human pro
39	32	80.0	196	7	ADJ92327	Adj92327 Mouse hai
40	32	80.0	267	2	AAW22303	Aaw22303 Rat CRIL
41	32	80.0	267	8	ADL27274	Adl27274 Amino aci
42	32	80.0	282	7	ADD30604	Add30604 Plant yie
43	32	80.0	282	8	ADI44109	Adi44109 Plant tra
44	32	80.0	282	8	ADI61449	Adi61449 A. thalia
45	32	80.0	282	8	ADO02425	Ado02425 Thalecres

## ALIGNMENTS

## RESULT 1

AAE08241

ID AAE08241 standard; peptide; 9 AA.

XX AAE08241;

XX

DT 01-NOV-2001 (first entry)

XX Human stratum corneum chymotrypsin enzyme peptide #6 (residues 4-12).

DE Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;

KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;

KW antisense therapy; malignant hyperplasia.

XX Homo sapiens.

OS WO200159158-A1.

XX 16-AUG-2001.

PD 07-FEB-2001; 2001WO-US003977.

PF 11-FEB-2000; 2000US-00502600.

XX (UTAR-) UNIV ARKANSAS.

XX O'brien TV;

PI WPI; 2001-514676/56.

DR Diagnosing cancer comprises detecting stratum corneum chymotrypsin

XX enzyme.

PT Claim 25; Page 103; 127pp; English.

PS The invention relates to diagnosing cancer especially ovarian cancer, by

XX screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are

CC considered to be an integral part of tumour growth and metastasis, and

CC therefore, markers indicative of their presence or absence are useful for

CC the diagnosis of cancer. The method is useful for diagnosing cancer.

CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used

CC to treat a cancer selected from ovarian, breast, lung, colon, prostate is

CC and other cancers in which SCCE is overexpressed. The present sequence is

CC human SCCE peptide

XX Sequence 9 AA;

SQ Query Match 100.0%; Score 40; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9  
| | | | |  
Db 1 SLLPLQL 9

RESULT 2  
ADR68797  
ID ADR68797 standard; peptide; 9 AA.  
XX AC ADR68797;  
XX DT 02-DEC-2004 (first entry)  
XX DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:36.  
XX KW serine protease; stratum corneum chymotrytic enzyme; SCCE;  
XX KW immune response; ovarian cancer; lung cancer; prostate cancer;  
XX KW pancreatic cancer; colon cancer.  
XX OS Homo sapiens.  
XX PN WO2004075723-A2.  
XX PD 10-SEP-2004.  
XX PF 20-FEB-2004; 2004WO-US005134.  
XX PR 21-FEB-2003; 2003US-00372521.  
XX PA (UYAR-) UNIV ARKANSAS.  
XX PI O'brien TJ, Cannon MJ, Santin A;  
XX DR WPI; 2004-653294/63.  
XX PT Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for  
PT vaccinating an individual against SCCE, and in monitoring and developing  
PT immunotherapies for ovarian and other malignancies.  
XX Claim 5; SEQ ID NO 36; 117pp; English.  
XX The invention relates to a novel method for vaccinating an individual  
CC against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating  
CC an individual with a SCCE peptide, which elicits an immune response in  
CC the individual. A peptide of the invention acts as a stratum corneum  
CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
CC an individual against SCCE, particularly an individual having, suspected  
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
CC an individual, such as ovarian, breast, lung, colon, prostate, or  
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
CC peptides are also useful in the monitoring and development of  
CC immunotherapies for ovarian and other malignancies. The present sequence  
CC represents a peptide fragment of serine protease SCCE (stratum corneum  
CC chymotrytic enzyme).  
XX SQ Sequence 9 AA;  
Query Match 100.0%; Score 40; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9  
| | | | |  
Db 1 SLLPLQL 9

RESULT 3  
ABG23378  
ID ABG23378 standard; protein; 136 AA.

XX ABG23378;  
XX AC 18-FEB-2002 (first entry)  
XX DT Novel human diagnostic protein #23369.  
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX DR WPI; 2001-639362/73.  
XX DR N-PSDB; AAS87565.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX Claim 20; SEQ ID NO 53737; 103pp; English.  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 136 AA;  
Query Match 100.0%; Score 40; DB 4; Length 136;  
Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9  
| | | | |  
Db 10 SLLPLQL 18

RESULT 4  
ADA05736  
ID ADA05736 standard; protein; 198 AA.  
XX AC ADA05736;

XX 06-NOV-2003 (first entry)

DE Human NOV18c protein SEQ ID NO:96.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytostatic; nootropic; neuroprotective;

KW antiparkinsonian; antilipidemic; gene therapy; human disease;

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;

PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI: 2003-381626/36.

DR N-PSDB; ADA05735.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX Claim 1; Page 170; 586pp; English.

PS The present invention describes NOVX proteins, where X can be 1 to 55

XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC binds to the polypeptide described above; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described

CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or

CC aberrant physiological interactions of the polypeptide; (11) a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating

CC the activity of the polypeptide described above; (13) methods of treating

CC or preventing a pathology associated with the above polypeptide in a

CC mammal; and (14) a method for producing the above polypeptide. NOVX

CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian

CC and antilipidemic activities, and can be used in gene therapy. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic

CC disorders such as diabetes or obesity, infections, cachexia, cancer,

CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation

CC probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence represents a human NOVX from the

CC present invention.

XX Sequence 198 AA;

SQ

Query Match 100.0%; Score 40; DB 6; Length 198;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9

Db |||||

4 SLLPLQL 12

RESULT 5

ADN62900

ID ADN62900 standard; protein; 198 AA.

AC ADN62900;

DT 01-JUL-2004 (first entry)

DE Human NOV18c.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;

KW Alzheimer's disease; Parkinson's disease; immune disorder;

KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;

KW wasting disorder.

OS Homo sapiens.

XX US2004038223-A1.

PD 26-FEB-2004.

PF 01-OCT-2002; 2002US-00262511.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJUV/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPVT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERR/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG W.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENNA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIP/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.  
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spyttek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX WPI; 2004-213931/20.  
 DR N-PSDB; ADN62899.  
 XX  
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 XX  
 PS Claim 1; SEQ ID NO 96; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX  
 SQ Sequence 198 AA;  
 Query Match 100.0%; Score 40; DB 8; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 Db 4 SLLPLQL 12  
 RESULT 6  
 ID ADA05732 standard; protein; 250 AA.  
 XX ADA05732;  
 AC ADA05732;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV18a protein SEQ ID NO:92.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003029424-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 02-OCT-2002; 2002WO-US031373.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.



12-OCT-2001; 2001US-0328849P.  
15-OCT-2001; 2001US-0329414P.  
17-OCT-2001; 2001US-0330142P.  
18-OCT-2001; 2001US-0330309P.  
22-OCT-2001; 2001US-0341058P.  
24-OCT-2001; 2001US-0339266P.  
24-OCT-2001; 2001US-0343629P.  
29-OCT-2001; 2001US-0349575P.  
29-OCT-2001; 2001US-0345357P.  
17-APR-2002; 2002US-0373260P.  
19-APR-2002; 2002US-0373815P.  
19-APR-2002; 2002US-0373817P.  
19-APR-2002; 2002US-0373826P.  
19-APR-2002; 2002US-0373884P.  
22-APR-2002; 2002US-0374977P.  
22-APR-2002; 2002US-0381037P.  
16-MAY-2002; 2002US-0381038P.  
16-MAY-2002; 2002US-0381042P.  
17-MAY-2002; 2002US-0381642P.  
28-MAY-2002; 2002US-0383656P.  
29-MAY-2002; 2002US-0383831P.  
25-JUN-2002; 2002US-0391335P.  
01-OCT-2002; 2002US-00262511.  
(CURA-) CURAGEN CORP.  
Smithson G, Millett I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
Patturajan M, Szyte KA, Edinger SR, Ellerman K, Malyankar UM;  
Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
WPI; 2003-381626/36.  
N-PSDB; ADA05731.  
New NOVX polypeptides and nucleic acids, useful for diagnosing,  
preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
pharmacogenomics.  
Claim 1; Page 169-170; 586pp; English.  
The present invention describes NOVX proteins, where X can be 1 to 55  
(e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
described above and a carrier; (2) a kit comprising, in one or more  
containers, the composition described above; (3) an isolated nucleic acid  
molecule which encodes a NOVX protein of the invention; (4) a vector  
comprising the nucleic acid molecule described above; (5) a cell  
comprising the above vector; (6) an antibody that immunospecifically  
binds to the polypeptide described above; (7) methods for determining the  
presence or amount of the above polypeptide or nucleic acid molecule in a  
sample; (8) methods for determining the presence of or predisposition to  
a disease associated with altered levels of expression of the above  
polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
method of identifying an agent that binds to the polypeptide described  
above; (10) a method for identifying a potential therapeutic agent for  
use in treating a pathology that is related to an aberrant expression or  
aberrant physiological interactions of the polypeptide; (11) a method of  
screening for a modulator of activity or of latency or predisposition to  
a pathology associated with the polypeptide; (12) a method for modulating  
the activity of the polypeptide described above; (13) methods of treating  
or preventing a pathology associated with the above polypeptide in a  
mammal; and (14) a method for producing the above polypeptide in a  
sequence have antidiabetic, anorectic, antibacterial, virucide,  
immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
and antilipidemic activities, and can be used in gene therapy. The  
polypeptide is useful in manufacturing a medicament for treating a  
syndrome associated with a human disease. The polypeptide or the nucleic  
acid molecule may be used to diagnose, treat or prevent metabolic  
disorders such as diabetes or obesity, infections, cachexia, cancer,  
neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX

SQ Sequence 250 AA;

Query Match 100.0%; Score 40; DB 6; Length 250;  
Blast Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLLLPQLIL 9  
|||||||  
Db 1 SLLLPQLIL 9

#### RESULT 7

ADN62896  
ID ADN62896 standard; protein; 250 AA.

XX

AC ADN62896;

XX 01-JUL-2004 (first entry)

XX Human NOV18a.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.

XX Homo sapiens.

XX US2004038223-A1.

XX 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.  
XX 02-OCT-2001; 2001US-0326483P.  
XX 05-OCT-2001; 2001US-0327435P.  
XX 05-OCT-2001; 2001US-0327449P.  
XX 09-OCT-2001; 2001US-0327917P.  
XX 09-OCT-2001; 2001US-0328029P.  
XX 09-OCT-2001; 2001US-0328044P.  
XX 12-OCT-2001; 2001US-0328056P.  
XX 15-OCT-2001; 2001US-0328849P.  
XX 17-OCT-2001; 2001US-0329414P.  
XX 18-OCT-2001; 2001US-0330142P.  
XX 22-OCT-2001; 2001US-0330309P.  
XX 24-OCT-2001; 2001US-0339266P.  
XX 24-OCT-2001; 2001US-0343629P.  
XX 29-OCT-2001; 2001US-0349575P.  
XX 01-NOV-2001; 2001US-0346357P.  
XX 17-APR-2002; 2002US-0373260P.  
XX 19-APR-2002; 2002US-0373815P.  
XX 19-APR-2002; 2002US-0373817P.  
XX 19-APR-2002; 2002US-0373826P.  
XX 19-APR-2002; 2002US-0373884P.  
XX 22-APR-2002; 2002US-0374977P.  
XX 16-MAY-2002; 2002US-0381037P.  
XX 16-MAY-2002; 2002US-0381038P.  
XX 17-MAY-2002; 2002US-0381642P.  
XX 28-MAY-2002; 2002US-0383656P.  
XX 29-MAY-2002; 2002US-0383831P.  
XX 25-JUN-2002; 2002US-0391335P.

(SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU//) KEKUDA R.  
PA (JUJJ//) JU J.  
PA (LILL//) LI L.  
PA (GUOX//) GUO X.  
PA (PATT//) PATTURAJAN M.  
PA (SPYT//) SPYTEK K A.  
PA (EDIN//) EDINGER S R.  
PA (ELLE//) ELLERMAN K.  
PA (MALY//) MALYANKAR U M.  
PA (ORTT//) ORT T.  
PA (GORM//) GORMAN L.  
PA (ZERH//) ZERHUSEN B D.  
PA (ANDE//) ANDERSON D W.  
PA (ZHON//) ZHONG M.  
PA (CATT//) CATTERTON E.  
PA (JIWW//) JI W.  
PA (MILL//) MILLER C E.  
PA (RAST//) RASTELLI L.  
PA (STON//) STONE D J.  
PA (PENA//) PENNA C E A.  
PA (SHEN//) SHENOY S G.  
PA (SHIM//) SHIMKETS R A.  
PA (ROTH//) ROTHENBERG M E.  
PA (LEAC//) LEACH M D.  
PA (AGEE//) AGEE M L.  
PA (BERG//) BERGHS C.  
PA (DIPI//) DIPIPPO V A.  
PA (EISE//) EISEN A.  
PA (GANG//) GANGOLLI E A.  
PA (RIEG//) RIEGER D K.  
PA (SPAD//) SPADERNA S K.  
XX  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
XX WPI; 2004-211931/20.  
DR N-PSDB; ADN62895.  
XX  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
XX Claim 1; SEQ ID NO 92; 395pp; English.  
PS  
XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disorders, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC haematopoietic disorders, and the various dyslipidaemias, immune disorders,  
CC Alzheimer's Disease, Parkinson's Disease, metabolic disorders, and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC

CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 250 AA;  
Query Match 100.0%; Score 40; DB 8; Length 250;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SLLPLQL 9  
Db 1 SLLPLQL 9  
RESULT 8  
AAR67888  
ID AAR67888 standard; protein; 253 AA.  
XX  
XX AAR67888;  
XX  
XX 25-MAR-2003 (revised)  
DT 09-AUG-1995 (first entry)  
XX  
XX Human stratum corneum chymotrophic recombinant enzyme (SCCE).  
DE Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
KW callosities; keratosis pilaris; ichthyoses; eczema.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9500651-A1.  
PN  
XX  
XX 05-JAN-1995.  
PD  
XX  
XX 20-JUN-1994; 94WO-IB000166.  
PF  
XX  
XX 18-JUN-1993; 93DK-00000725.  
PR  
XX  
XX (SYMB-) SYMBICOM AB.  
PA  
XX  
XX Egelrud T, Hansson L;  
PI  
XX  
XX WPI; 1995-052088/07.  
DR  
XX  
XX N-PSDB; AAQ81203.  
DR  
PT Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
PT related vectors, transformed cells and polypeptides, useful for treating  
PT skin disorders, e.g. acne or psoriasis, and for identification of  
PT specific inhibitors.  
XX  
XX Disclosure; Page 97; 137pp; English.  
PS  
XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
CC and skin care products, especially to treat and prevent acne, xeroderma,  
CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
CC recombinantly following mammal, insect, plant, or microorganism  
CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX  
XX Sequence 253 AA;  
Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12  
RESULT 9  
AAW05383

ID AAW05383 standard; protein; 253 AA.  
 AC AAW05383;  
 XX  
 DT 31-DEC-1996 (first entry)  
 XX  
 DE Human amyloid precursor protein protease.  
 XX  
 KW Amyloid precursor protein protease; Alzheimer's disease; diagnosis;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9631122-A1.  
 XX  
 PD 10-OCT-1996.  
 XX  
 PF 02-APR-1996; 96WO-US004294.  
 XX  
 PR 04-APR-1995; 95US-00416257.  
 XX  
 PA (ELIL ) LILLY & CO ELI.  
 XX  
 PI Dixon EP, Johnstone EM, Little SP;  
 XX  
 DR WPI; 1996-464694/46.  
 DR N-PSDB; AAT39783.  
 XX  
 PT New isolated human amyloid precursor protein protease - used to develop  
 PT prods. for the treatment or diagnosis of associated conditions, esp.  
 PT Alzheimer's disease.  
 XX  
 PS Claim 1; Page 44-45; 55pp; English.  
 XX  
 CC Human amyloid precursor protein protease (AAW05383) is involved in the  
 CC processing or clearance of amyloid precursor protein to form beta-amyloid  
 CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)  
 CC obtd. from a human lung library. Recombinant protease can be produced in  
 CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic  
 CC (partic. AV-120 host cells. It is used to develop products for the design  
 CC and testing of cpds. useful for treating or preventing conditions  
 CC associated with beta-amyloid peptide, esp. Alzheimer's disease  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 DB 4 SLLPLQL 12  
 RESULT 10  
 ABB84421  
 ID ABB84421 standard; peptide; 253 AA.  
 XX  
 AC ABB84421;  
 XX  
 DT 08-NOV-2002 (first entry)  
 XX  
 DE Human SCCE protein N-terminal fragment SEQ ID 48.  
 XX  
 KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
 KW pruritus; atopic dermatitis; eczema; acne; itch; KUK7; enzyme.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200262135-A2.  
 XX

PD 15-AUG-2002.  
 XX  
 PF 08-FEB-2002; 2002WO-IB001300.  
 XX  
 PR 09-FEB-2001; 2001CA-02332655.  
 PR 09-FEB-2001; 2001DK-00000218.  
 XX  
 PA (EGEL/) EGELRUD T.  
 PA (HANS/) HANSSON L.  
 XX  
 PI Egelrud T, Hansson L;  
 XX  
 DR WPI; 2002-643380/69.  
 XX  
 PT Transgenic mammal or its embryo useful as model for human disease, has  
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
 PT enzyme operably linked to promoter that drives its expression in skin.  
 XX  
 PS Example 6; Page 37; 74pp; English.  
 XX  
 CC This invention describes a novel non-human transgenic mammal or mammalian  
 CC embryo having integrated within its genome, a heterologous nucleotide  
 CC sequence comprising at least a significant part of a nucleotide sequence  
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
 CC operably linked to a promoter that drives expression of heterologous scce  
 CC or its variant in skin. The product of the invention is useful as a model  
 CC for the study of disease with the aim of improving treatment, to relieve  
 CC or ameliorate a pathological condition, for development or testing of a  
 CC cosmetic or a pharmaceutical formulation, and for the development of a  
 CC diagnostic method. It can also be used as a model for a skin disease or  
 CC skin cancer. The invention is also useful for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of an  
 CC abnormal or unwanted phenotype, and for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of  
 CC inflammatory skin diseases selected from diseases consisting of epidermal  
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
 CC as a model for further studies of itch mechanisms and the testing of  
 CC potential compounds and compositions for relieve of various skin diseases  
 CC where itch is a component. This sequence represents the N-terminal  
 CC fragment of the human stratum corneum chymotryptic enzyme, SCCE  
 CC synonymous with human kallikrein 7 (KLK7), used in the development of the  
 CC transgenic mammals described in the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 DB 4 SLLPLQL 12  
 RESULT 11  
 ABB84406  
 ID ABB84406 standard; protein; 253 AA.  
 XX  
 AC ABB84406;  
 XX  
 DT 08-NOV-2002 (first entry)  
 XX  
 DE Human SCCE protein.  
 XX  
 KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
 KW pruritus; atopic dermatitis; eczema; acne; itch; KUK7; enzyme.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200262135-A2.  
 XX 15-AUG-2002.  
 XX 08-FEB-2002; 2002WO-IB001300.  
 XX 09-FEB-2001; 2001CA-02332655.  
 PR 09-FEB-2001; 2001DK-00000218.  
 XX (EGEL/) EGELRUD T.  
 PA (HANS/) HANSSON L.  
 XX  
 XX Egelrud T, Hansson L;  
 XX  
 XX WPI; 2002-643380/69.  
 DR N-PSDB; ABQ76226.  
 XX  
 XX Transgenic mammal or its embryo useful as model for human disease, has  
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
 PT enzyme operably linked to promoter that drives its expression in skin.  
 XX  
 PS Claim 10; Page 58-59; 74pp; English.  
 XX  
 CC This invention describes a novel non-human transgenic mammal or mammalian  
 CC embryo having integrated within its genome, a heterologous nucleotide  
 CC sequence comprising at least a significant part of a nucleotide sequence  
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
 CC operably linked to a promoter that drives expression of heterologous scce  
 CC or its variant in skin. The product of the invention is useful as a model  
 CC for the study of disease with the aim of improving treatment, to relieve  
 CC or ameliorate a pathogenic condition, for development or testing of a  
 CC cosmetic or a pharmaceutical formulation, and for the development of a  
 CC diagnostic method. It can also be used as a model for a skin disease or  
 CC skin cancer. The invention is also useful for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of an  
 CC abnormal or unwanted phenotype, and for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of  
 CC inflammatory skin diseases selected from diseases consisting of epidermal  
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
 CC as a model for further studies of itch mechanisms and the testing of  
 CC potential compounds and compositions for relieve of various skin diseases  
 CC where itch is a component. This sequence represents the human stratum  
 CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymous  
 CC with human kallikrein 7 (KLK7) and is used in the development of the  
 CC transgenic mammals described in the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 Db |||||  
 4 SLLPLQL 12  
 RESULT 12  
 AAU82740  
 ID AAU82740 standard; protein; 253 AA.  
 XX  
 AC AAU82740;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE Amino acid sequence of novel human protease #39.  
 XX  
 KW Human; protease; cancer; immune-related disorder; cardiovascular disease;  
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;

KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
 KW ocular disease; cytostatic; enzyme.  
 OS Homo sapiens.  
 XX  
 PN WO200200860-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 26-JUN-2001; 2001WO-US020171.  
 XX  
 PR 26-JUN-2000; 2000US-0214047P.  
 XX  
 PA (SUGE-) SUGEN INC.  
 XX  
 PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
 PI Charyczak G;  
 XX  
 XX WPI; 2002-139913/18.  
 DR N-PSDB; ABK31782.  
 XX  
 PT Nucleic acids encoding novel human proteases, useful for useful for  
 PT treating diseases and disorders such as cancers, immune-related diseases  
 PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory  
 PT disorders.  
 XX  
 XX Claim 6; Fig 2N; 313pp; English.  
 CC  
 CC The present invention relates to the isolation of novel human proteases,  
 CC and the nucleic acids encoding them. The sequences of the invention are  
 CC useful for treating diseases and disorders such as cancers (e.g. breast,  
 CC colon, lung), immune-related diseases and disorders (e.g. inflammatory  
 CC diseases and asthma), cardiovascular diseases (e.g. restenosis and  
 CC coronary thrombosis), brain or neuronal-associated diseases, metabolic  
 CC disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.  
 CC rheumatoid arthritis and psoriasis), central or peripheral nervous system  
 CC diseases, migraines, pain, sexual dysfunction, mood disorders, attention  
 CC disorders, neurological disorders, hypotension, hypertension, psychotic  
 CC disease) and dyskinesias. The nucleic acids and polypeptides are also  
 CC useful for treating viral infections caused by human immunodeficiency  
 CC virus (HIV), and non-viral infections such as ocular disease (e.g.  
 CC glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel  
 CC human proteases of the invention  
 XX  
 SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 Db |||||  
 4 SLLPLQL 12  
 RESULT 13  
 ABU07440  
 ID ABU07440 standard; protein; 253 AA.  
 XX  
 AC ABU07440;  
 XX  
 DT 28-JAN-2003 (first entry)  
 XX  
 DE Protein differentially regulated in prostate cancer #43.  
 XX  
 KW Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200281638-A2.

XX PD 17-OCT-2002.  
 XX PF 08-APR-2002; 2002WO-US010824.  
 XX PR 06-APR-2001; 2001US-0281731P.  
 XX PR 06-APR-2001; 2001US-0281732P.  
 XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 XX PI Sun Z, Jay G;  
 XX DR N-PSDB; ABX10343.  
 XX PT Novel genes which are differentially regulated in prostate cancer, useful  
 XX PT for diagnosing prostate cancer in prostate tissue sample and assessing  
 XX PT therapeutic or preventive intervention in prostate cancer patients.  
 XX PS Claim 1; Page 293-294; 416pp; English.  
 XX CC The invention describes genes (I) which are differentially regulated in  
 XX CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 XX CC sample comprising prostate tissue, which involves determining the number  
 XX CC of target genes which are differentially-regulated in the sample, where  
 XX CC the number is indicative of the probability that the sample comprises  
 XX CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 XX CC intervention in a subject having a prostate cancer, which involves  
 XX CC determining the expression levels in a sample comprising prostate tissue  
 XX CC of target genes which are differentially-regulated in prostate cancer.  
 XX CC Preferably, the expression levels of at least 10 genes are determined.  
 XX CC (I) is also useful for identifying agents that modulate a biological  
 XX CC activity of a polypeptide differentially-regulated in prostate cancer  
 XX CC cells, which involves contacting a polypeptide differentially-regulated  
 XX CC in prostate cancer cells with a test agent under conditions effective for  
 XX CC the test agent to modulate a biological activity of the polypeptide, and  
 XX CC determining whether the test agent modulates the biological activity. (I)  
 XX CC is useful as molecular markers, as drug targets, and for detecting,  
 XX CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
 XX CC preventing or treating, determining predisposition to diseases and  
 XX CC conditions especially relating to prostate cancer. (I) and its expression  
 XX CC products are used in the diagnostic test to assay for presence of cancer  
 XX CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 XX CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 XX CC of cancer, its stage of development, the nature of genetic defect, etc.  
 XX CC The polypeptide encoded by (I) can be used as target for therapy or drug  
 XX CC discovery. (I) can also be used for expressing the polypeptide and thus  
 XX CC for searching specific binding partners of the polypeptide. (I) is useful  
 XX CC in therapeutic applications to treat prostate cancer. The identification  
 XX CC of specific genes, and groups of genes, expressed in pathways  
 XX CC physiologically relevant to prostate cancer permits the definition of  
 XX CC functional and disease pathways and the delineation of targets in these  
 XX CC pathways which are useful in diagnostic, therapeutic, and clinical  
 XX CC applications. This is the amino acid sequence of a protein differentially  
 XX CC regulated in prostate cancer  
 XX SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 Db |||||  
 4 SLLPLQL 12  
 RESULT 14  
 ABU07471  
 ID ABU07471 standard; protein; 253 AA.  
 XX AC ABU07471;  
 XX

DT 28-JAN-2003 (first entry)  
 XX DE Protein differentially regulated in prostate cancer #74.  
 XX KW Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX OS Homo sapiens.  
 XX PN WO200281638-A2.  
 XX PD 17-OCT-2002.  
 XX PF 08-APR-2002; 2002WO-US010824.  
 XX PR 06-APR-2001; 2001US-0281731P.  
 XX PR 06-APR-2001; 2001US-0281732P.  
 XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 XX PI Sun Z, Jay G;  
 XX DR WPI; 2003-058520/05.  
 DR N-PSDB; ABX10375.  
 XX PT Novel genes which are differentially regulated in prostate cancer, useful  
 XX PT for diagnosing prostate cancer in prostate tissue sample and assessing  
 XX PT therapeutic or preventive intervention in prostate cancer patients.  
 XX PS Claim 1; Page 351; 416pp; English.  
 XX CC The invention describes genes (I) which are differentially regulated in  
 XX CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 XX CC sample comprising prostate tissue, which involves determining the number  
 XX CC of target genes which are differentially-regulated in the sample, where  
 XX CC the number is indicative of the probability that the sample comprises  
 XX CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 XX CC intervention in a subject having a prostate cancer, which involves  
 XX CC determining the expression levels in a sample comprising prostate tissue  
 XX CC of target genes which are differentially-regulated in prostate cancer.  
 XX CC Preferably, the expression levels of at least 10 genes are determined.  
 XX CC (I) is also useful for identifying agents that modulate a biological  
 XX CC activity of a polypeptide differentially-regulated in prostate cancer  
 XX CC cells, which involves contacting a polypeptide differentially-regulated  
 XX CC in prostate cancer cells with a test agent under conditions effective for  
 XX CC the test agent to modulate a biological activity of the polypeptide, and  
 XX CC determining whether the test agent modulates the biological activity. (I)  
 XX CC is useful as molecular markers, as drug targets, and for detecting,  
 XX CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
 XX CC preventing or treating, determining predisposition to diseases and  
 XX CC conditions especially relating to prostate cancer. (I) and its expression  
 XX CC products are used in the diagnostic test to assay for presence of cancer  
 XX CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 XX CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 XX CC of cancer, its stage of development, the nature of genetic defect, etc.  
 XX CC The polypeptide encoded by (I) can be used as target for therapy or drug  
 XX CC discovery. (I) can also be used for expressing the polypeptide and thus  
 XX CC for searching specific binding partners of the polypeptide. (I) is useful  
 XX CC in therapeutic applications to treat prostate cancer. The identification  
 XX CC of specific genes, and groups of genes, expressed in pathways  
 XX CC physiologically relevant to prostate cancer permits the definition of  
 XX CC functional and disease pathways and the delineation of targets in these  
 XX CC pathways which are useful in diagnostic, therapeutic, and clinical  
 XX CC applications. This is the amino acid sequence of a protein differentially  
 XX CC regulated in prostate cancer  
 XX SQ Sequence 253 AA;  
 Query Match 100.0%; Score 40; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 40; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9  
| | | | |  
Db 4 SLLPLQL 12

RESULT 15  
ABR58471  
ID ABR58471 standard; protein; 253 AA.  
XX  
AC ABR58471;  
XX  
DT 07-JUL-2003 (first entry)  
XX  
DE Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.  
XX  
KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029468-A1.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031467.  
XX  
PR 02-OCT-2001; 2001US-0327135P.  
PR 30-MAY-2002; 2002US-0384531P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Algate PA, Mannion J;  
XX  
XX WPI; 2003-372001/35.  
XX  
PT New polynucleotide and polypeptide useful for diagnosing and/or treating  
PT cancer, particularly ovarian cancer, and as a vaccine.  
XX  
PS Claim 2; Page 157-158; 169pp; English.  
XX  
CC The invention relates to a novel isolated polynucleotide. The  
CC polynucleotides of the invention have cytostatic activity, and may have a  
CC use in gene therapy, and in a vaccine. The composition and methods are  
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.  
CC The composition may also be used as a vaccine to prevent cancer. The  
CC present sequence is used in the exemplification of the invention  
XX  
SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9  
| | | | |  
Db 4 SLLPLQL 12

Search completed: July 13, 2005, 17:19:54  
Job time : 77.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083A-36

Perfect score: 40

Sequence: 1 SLLPLQL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	253	2 A53968	serine proteinase
2	38	95.0	146	2 H75201	hypothetical prote
3	33	82.5	370	2 AB3334	daunorubicin resis
4	32	80.0	218	2 T02912	probable transcrip
5	32	80.0	264	2 C97402	probable acyltrans
6	32	80.0	266	2 AC2620	1-acyl-sn-glycerol
7	32	80.0	266	2 JC7300	tax-responsive ele
8	32	80.0	267	2 JC4857	hepatocarcinogenes
9	32	80.0	282	2 B85327	probable transcrip
10	32	80.0	494	2 C49349	probable succinogl
11	32	80.0	494	2 B95976	probable transport
12	32	80.0	582	2 S40176	ExoT protein - Rhi
13	32	80.0	622	2 S61692	probable membrane
14	31	77.5	196	2 G65039	hypothetical prote
15	31	77.5	303	2 C84914	hypothetical prote
16	31	77.5	354	2 T48649	glycerol-3-phospha
17	31	77.5	372	2 A98157	probable permease
18	31	77.5	372	2 AH3130	ABC transporter, m
19	31	77.5	392	2 D83934	hypothetical prote
20	31	77.5	395	2 D69779	antibiotic resista
21	31	77.5	398	2 C91063	hypothetical prote
22	31	77.5	401	2 B90120	SNF1-related prote
23	31	77.5	413	2 AC0834	probable membrane
24	31	77.5	470	2 AG0083	hypothetical prote
25	31	77.5	475	1 A69149	O-antigen transpor
26	31	77.5	754	2 AE0614	probable competenc
27	31	77.5	783	2 A46136	myosin-heavy-chain
28	31	77.5	913	2 AC2445	hypothetical prote
29	30	75.0	155	2 I45913	interleukin-2 prec

30	30	75.0	155	2 S38662	interleukin-2 - go
31	30	75.0	155	2 S11498	interleukin-2 prec
32	30	75.0	168	2 E96979	probable membrane
33	30	75.0	189	2 T43766	hypothetical prote
34	30	75.0	190	2 T39622	probable phosphate
35	30	75.0	208	2 T17092	NADH2 dehydrogenas
36	30	75.0	214	1 G69798	hypothetical prote
37	30	75.0	223	2 F83598	cell division prot
38	30	75.0	277	2 H84314	cytochrome aa3 con
39	30	75.0	289	2 T43663	probable high affi
40	30	75.0	333	1 HLHUCB	T-cell surface gly
41	30	75.0	349	2 B82656	conserved hypothet
42	30	75.0	369	2 S33603	surfactant protein
43	30	75.0	371	1 UN0450	conglutinin precur
44	30	75.0	371	2 I45878	conglutinin - bovi
45	30	75.0	375	2 S47704	hypothetical 41.1K

#### ALIGNMENTS

##### RESULT 1

A53968

serine proteinase SCCE precursor - human

N:Alternate names: stratum corneum chymotryptic enzyme

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A53968

R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A>Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A:Reference number: A53968; MUID:94308225; PMID:8034709

A:Accession: A53968

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <HAN>

A:Cross-references: UNIPROT:P49862; GB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G532504

C:Genetics:

A:Gene: GDB:PRSS6; SCCE

A:Cross-references: GDB:377730

A:Map position: 7q35-7q35

C:Superfamily: trypsin; trypsin homology

F/30-245/Domain: trypsin homology <TRY>

Query Match Similarity 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9

Db 4 SLLPLQL 12

##### RESULT 2

H75201

hypothetical protein PAB0088 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C:Accession: H75201

R:Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc

A:Reference number: A75001

A:Accession: H75201

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-146 <KAW>

A:Cross-references: UNIPROT:Q9V2D5; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CAB4906;

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0088

C:Superfamily: Pyrococcus abyssi hypothetical protein PAB0088

```
Query Match      95.0%; Score 38; DB 2; Length 146;
Best Local Similarity 88.9%; Pred. No. 2.3;
Matches      8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLLPLQIL 9
Db      66 SLLPLQII 74

RESULT 3
AB3334
daunorubicin resistance transmembrane protein [imported] - Brucella melitensis (strain 1
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AB3334
R:DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltman, E.; Selkov, E.; Elizer, P.H.; Hagiub, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-370 <KUR>
A:Cross-references: UNIPROT:Q8YHY9; UNIPROT:Q8FZX1; GB:AE008917; PIDN:AAL51837.1; PID:gl
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0656
A:Map position: 1

Query Match      82.5%; Score 33; DB 2; Length 370;
Best Local Similarity 75.0%; Pred. No. 59;
Matches      6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LLLPLQIL 9
Db      293 ILLPLQVL 300

RESULT 4
T02912
probable transcription factor MYB41 [similarity] - Arabidopsis thaliana (fragment)
N:Alternate names: protein T13J8.220
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Aug-2004
C:Accession: T02912; T51654
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schuel
submitted to the Protein Sequence Database, February 1999
A:Reference number: T214766
A:Accession: T02912
A:Molecule type: DNA
A:Residues: 1-218 <BEV>
A:Cross-references: UNIPROT:Q9ZTE3; EMBL:AL035524
A:Experimental source: cultivar Columbia; BAC clone T13J8
R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; J
; Paz-Ares, J.; Weisshaar, B.
Plant J. 16, 263-276, 1998
A:Title: Towards functional characterisation of the members of the R2R3-MYB gene from A
A:Reference number: Z14349; MUID:9839469; PMID:9839469
A:Accession: T51654
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-218 <KRA>
A:Cross-references: EMBL:AF062882; PIDN:AAC83604.1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: MYB41
A:Map position: 4
A:Note: T13J8.220
A:Note: intron positions not resolved
C:Superfamily: myb DNA-binding repeat homology
C:Keywords: transcription factor
F:1-48/Domain: myb DNA-binding repeat homology (fragment) <MYB>
```

```
Query Match      80.0%; Score 32; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 54;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLLPLQ 7
Db      104 SLLPLQ 110

RESULT 5
C97402
probable acyltransferase (AF232919) [imported] - Agrobacterium tumefaciens (strain C58, (
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 16-Aug-2004
C:Accession: C97402
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C97402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: UNIPROT:Q8UIE2; GB:AE007869; PIDN:AAK86172.1; PID:gl5155265; GSPDB:G
C:Genetics:
A:Gene: AGR_C_621
A:Map position: circular chromosome
C:Superfamily: 1-acyl-sn-glycerol-3-phosphate acyltransferase

Query Match      80.0%; Score 32; DB 2; Length 264;
Best Local Similarity 75.0%; Pred. No. 66;
Matches      6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LLLPLQIL 9
Db      19 ILLPLQLL 26

RESULT 6
AC2620
1-acyl-sn-glycerol-3-phosphate acyltransferase plsc [imported] - Agrobacterium tumefaciar
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 16-Aug-2004
C:Accession: AC2620
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC2620
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: UNIPROT:Q8UIE2; GB:AE008688; PIDN:AAL41377.1; PID:gl7738693; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: plsc
A:Map position: circular chromosome
C:Superfamily: 1-acyl-sn-glycerol-3-phosphate acyltransferase

Query Match      80.0%; Score 32; DB 2; Length 264;
Best Local Similarity 75.0%; Pred. No. 66;
Matches      6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LLLPLQIL 9
Db      19 ILLPLQLL 26
```



## RESULT 7

JC7300  
tax-responsive element-binding protein 5 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: JC7300  
R:Masaki, T.; Noguchi, H.; Kobayashi, M.; Yoshida, M.; Takamatsu, K.  
DNA Res. 7, 187-193, 2000  
A:Title: Isolation and characterization of the gene encoding mouse tax-responsive element  
A:Reference number: JC7300  
A:Accession: JC7300  
A:Molecule type: mRNA  
A:Residues: 1-266 <MAS>  
A:Cross-references: UNIPROT:Q9ESS3; DBJ:AB036745  
C:Genetics:  
A:Gene: treb5  
C:Superfamily: X box-binding protein 1; fos/jun DNA-binding domain homology  
C:Keywords: leucine zipper; transcription factor

Query Match 80.0%; Score 32; DB 2; Length 266;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 LLPLQIL 9  
DB 187 LLPLQIL 193

RESULT 8  
JC4857  
hepatocarcinogenesis-related transcription factor - rat

C:Species: Rattus norvegicus (Norway rat)  
C:Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 09-Jul-2004  
C:Accession: JC4857  
R:Kishimoto, T.; Kokura, K.; Kumagai, Y.; Makino, Y.; Tamura, T.  
Biochem. Biophys. Res. Commun. 224, 746-751, 1996  
A:Title: Hrf: A b-zip transcription factor that is closely related to the human XBP/TRE  
A:Reference number: JC4857  
A:Accession: JC4857  
A:Molecule type: mRNA  
A:Residues: 1-267 <KIS>  
A:Cross-references: UNIPROT:Q9R1S4  
C:Comment: This is a basic-leucine zipper type transcription factor involved in hepatocarcinogenesis  
C:Genetics:  
A:Gene: htf  
C:Superfamily: X box-binding protein 1; fos/jun DNA-binding domain homology  
C:Keywords: leucine zipper; transcription factor  
F:58-98/Domain: fos/jun DNA-binding domain homology <FJD>  
F:78-126/Region: leucine zipper motif

Query Match 80.0%; Score 32; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 LLPLQIL 9  
DB 188 LLPLQIL 194

RESULT 9  
B85327  
probable transcription factor MYB41 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: B85327  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: B85327  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-282 <STO>

A:Cross-references: UNIPROT:Q9MOJ5; GB:NC\_001268; NID:97269665; PIDN:CAB79613.1; GSPDB:G1C7300

C:Genetics:  
A:Gene: AT4g28110  
A:Map position: 4  
C:Superfamily: barley myb-related protein 3; myb DNA-binding repeat homology

Query Match 80.0%; Score 32; DB 2; Length 282;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQ 7  
DB 168 SLLPLQ 174

## RESULT 10

C49349  
probable succinoglycan transport protein ExoT - Rhizobium meliloti  
C:Species: Rhizobium meliloti  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C:Accession: C49349  
R:Glucksmann, M.A.; Reuber, T.L.; Walker, G.C.  
J. Bacteriol. 175, 7045-7055, 1993

A:Title: Genes needed for the modification, polymerization, export, and processing of sucrose  
A:Reference number: A49349; MUID:94042870; PMID:8226646  
A:Accession: C49349  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-494 <GLU>  
A:Cross-references: UNIPROT:P33699; GB:L20758; NID:g393240; PIDN:AAAI6050.1; PID:g393249  
C:Superfamily: hypothetical protein b2046  
C:Keywords: transmembrane protein

Query Match 80.0%; Score 32; DB 2; Length 494;  
Best Local Similarity 75.0%; Pred. No. 1-2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQI 8  
DB 122 SLLPLQL 129

## RESULT 11

B95976  
probable transport protein, similar to Wzx exoT [imported] - Sinorhizobium meliloti (strain 1021)  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: B95976  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernandez, T.M.; Weidner, S.; U.S.A. 98, 9889-9894, 2001

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endorhizobium strain 1021  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: B95976  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-494 <KUR>  
A:Cross-references: UNIPROT:P33699; GB:AL591985; PIDN:CAC49474.1; PID:g15140960; GSPDB:G1C7300  
A:Experimental source: strain 1021, megaplasmid pSymB  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, G.; Leblond, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
A:Genetics:  
C:Gene: exoT; SMB20950  
A:Genome: plasmid  
C:Superfamily: hypothetical protein b2046

```
Query Match      80.0%; Score 32; DB 2; Length 494;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQI 8
    |||:|:|
Db 122 SLLPLQL 129

RESULT 12
S40176
ExoT protein - Rhizobium meliloti
C:Species: Rhizobium meliloti
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 19-May-2000
C:Accession: S40176
R:Becker, A.; Kleickmann, A.; Kuester, H.; Keller, M.; Arnold, W.; Puhler, A.
submitted to the EMBL Data Library, April 1993
A:Description: Analysis of the Rhizobium meliloti genes exoU, exoV, exoW, exoT and exoI
ses.
A:Reference number: S40173
A:Accession: S40176
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <BEC>
A:Cross-references: EMBL:Z22646
C:Superfamily: hypothetical protein b2046

Query Match      80.0%; Score 32; DB 2; Length 582;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQI 8
    |||:|:|
Db 122 SLLPLQL 129

RESULT 13
S61692
probable membrane protein YOR137c - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O3329; hypothetical protein YOR3329c
C:Species: Saccharomyces cerevisiae
C>Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C:Accession: S61692; S67022
R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia
submitted to the EMBL Data Library, December 1995
A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
A:Reference number: S61643
A:Accession: S61692
A:Molecule type: DNA
A:Residues: 1-622 <BEN>
A:Cross-references: UNIPROT:Q12212; EMBL:X94335; NID:g1262139; PID:e217839; PID:g1164980
R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Faces, V.; Ansoerge, W.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66965
A:Accession: S67022
A:Molecule type: DNA
A:Residues: 1-622 <VOS>
A:Cross-references: EMBL:Z75045; NID:g1420348; PID:e252028; PID:g1420349; MIPS:YOR137c
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0005663
A:Map position: 15R
C:Keywords: transmembrane protein
F:11-27/Domain: transmembrane #status predicted <TMM>

Query Match      80.0%; Score 32; DB 2; Length 622;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLLPLQI 9
    |||:|:|
Db 141 SALLPLQVL 149
```

```
RESULT 14
G65039
hypothetical protein b2612 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: G65039
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97428617; PMID:9278503
A:Accession: G65039
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-196 <BLAT>
A:Cross-references: GB:AE000347; GB:U00096; NID:g2367142; PIDN:AAC75661.1; PID:g1788965;
A:Experimental source: strain K-12, substrain MGL655

Query Match      77.5%; Score 31; DB 2; Length 196;
Best Local Similarity 87.5%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LLLPLQIL 9
    |||:|:|
Db 123 LLAPLQIL 130

RESULT 15
C84914
hypothetical protein At2g47360 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84914
R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; A
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84914
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-303 <STO>
A:Cross-references: UNIPROT:O22910; GB:AE002093; NID:g2275214; PIDN:AAB63836.1; GSPDB:GN
C:Genetics:
A:Gene: At2g47360
A:Map position: 2

Query Match      77.5%; Score 31; DB 2; Length 303;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLLPLQIL 9
    |||:|:|
Db 29 SLLPLSFL 37

Search completed: July 13, 2005, 17:31:31
Job time : 14.5556 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-36

Perfect score: 40

Sequence: 1 SLLPLQL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	40	100.0	66	2 Q6DTY1
2	40	100.0	253	1 KLK7_HUMAN
3	40	100.0	253	2 Q8N5N9
4	38	95.0	146	2 Q9V2D5
5	37	92.5	208	2 Q21527
6	34	85.0	73	2 Q8VCA9
7	34	85.0	104	2 Q8R5D6
8	34	85.0	138	2 Q6PKZ2
9	34	85.0	201	1 FK11_MOUSE
10	34	85.0	361	1 IHA_TRIVU
11	34	85.0	373	2 Q749Q7
12	33	82.5	148	2 Q6N500
13	33	82.5	302	2 Q6GNT6
14	33	82.5	339	1 YG47_XANAC
15	33	82.5	370	2 Q8YHY9
16	33	82.5	370	2 Q8FZX1
17	33	82.5	379	2 Q7U9J6
18	33	82.5	381	2 Q7V4U4
19	33	82.5	392	2 Q732B9
20	33	82.5	392	2 Q819M8
21	33	82.5	393	2 Q635X8
22	33	82.5	393	2 Q6HEL6
23	33	82.5	394	2 Q6HU78
24	33	82.5	1017	1 EM11_MOUSE
25	32	80.0	68	2 Q6KG65
26	32	80.0	72	2 Q8VBB6
27	32	80.0	78	2 Q7WXY8
28	32	80.0	102	2 Q8BRV0
29	32	80.0	171	2 Q7U5R4
30	32	80.0	196	2 Q8D226
31	32	80.0	205	2 Q653Z3

32	32	80.0	218	2	Q9ZTE3	Q9ZTE3 arabidopsis
33	32	80.0	230	2	Q876R3	Q876R3 emericella
34	32	80.0	256	2	Q922G5	Q922G5 mus musculus
35	32	80.0	264	2	Q8UIE2	Q8UIE2 agrobacteri
36	32	80.0	266	2	Q9ESS3	Q9ESS3 mus musculus
37	32	80.0	267	2	Q9R1S4	Q9R1S4 rattus norv
38	32	80.0	267	2	Q35426	Q35426 mus musculus
39	32	80.0	282	2	Q9M0J5	Q9M0J5 arabidopsis
40	32	80.0	404	2	Q86166	Q86166 porphyromon
41	32	80.0	435	2	Q7MVD2	Q7MVD2 porphyromon
42	32	80.0	494	1	EXOT_RHIME	P33699 rhizobium m
43	32	80.0	559	2	Q821I4	Q821I4 streptomyce
44	32	80.0	622	2	Q12212	Q12212 saccharomyc
45	32	80.0	662	2	Q969D0	Q969D0 giardia lam

#### ALIGNMENTS

##### RESULT 1

Q6DTY1 PRELIMINARY; PRT; 66 AA.  
AC Q6DTY1  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Kallikrein 7 splice variant 3.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Kishi T., Michael I.P., Diamandis E.P.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY646152; AAT66047.1; -  
SQ SEQUENCE 66 AA; 7171 MW; 82E1C392BC8222FDB CRC64;

Query Match 100.0%; Score 40; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12

##### RESULT 2

KLK7\_HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (HSCCE).  
GN Name=KLK7; Synonyms=PRSS6, SCCE;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
RX MEDLINE=94308225; PubMed=8034709;  
RA Hansson L., Stromqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;  
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase."  
RL J. Biol. Chem. 269:19420-19426(1994).  
RN [2]  
RP SEQUENCE FROM N.A.

RC TISSUE=Keratinocytes;  
RA Yousef G.M., Scorilas A., Diamandis E.P.;  
RT "Molecular characterization, mapping and tissue expression of the  
RL human stratum corneum chymotryptic enzyme gene.";   
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
RA Moss P., Paepel B., Wang K.;  
RA "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";   
RT Gene 257:119-130(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX Hanson L., Beckman A., Ny A., Edlund M., Edholm E., Tornell J.,  
RA Wallbrandt P., Egelrud T.;  
RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
RL mice; a model for chronic itchy dermatitis.";   
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE=95314630; PubMed=7794273;  
RA Skytt A., Stroemqvist M., Egelrud T.;  
RT "Primary substrate specificity of recombinant human stratum corneum  
RL chymotryptic enzyme.";   
RL Biochem. Biophys. Res. Commun. 211:586-589(1995).  
CC -I- FUNCTION: May catalyze the degradation of intercellular cohesive  
CC structures in the cornified layer of the skin in the continuous  
CC shedding of cells from the skin surface. Specific for amino acid  
CC residues with aromatic side chains in the P1 position. SCCE  
CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
CC Phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
CC activation of precursors to inflammatory cytokines.  
CC -I- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
CC expressed by keratinocytes in the epidermis. Very low levels are  
CC also seen in the brain and kidney.  
CC -I- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L33404; AAC37551.1; -;  
CC EMBL; AF166330; AAD49718.1; -;  
CC EMBL; AF243527; AAG33360.1; -;  
CC EMBL; AF325283; AAK69624.1; -;  
CC F01; A53968; A53968.  
CC HSP; P00760; IE2X.  
CC MEROPS; S01.300; -;  
CC Genew; HGNC:6368; KLK7.  
CC H-invDB; HIX0015373; -;  
CC MIM; 604438; -;  
CC GO; GO:0008236; F:serine-type peptidase activity; TAS.  
CC GO; GO:0008544; P:epidermal differentiation; TAS.  
CC InterPro; IPR009003; Pept\_Ser\_Cys.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; Trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; Tryp\_SPC; 1.  
CC PROSITE; PS0240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
KW Signal; Zymogen.  
FT SIGNAL 1 22  
FT PROPEP 23 29 Activation peptide.

FT CHAIN 30 253 Kallikrein 7.  
FT ACT\_SITE 70 70 Charge relay system (By similarity).  
FT ACT\_SITE 112 112 Charge relay system (By similarity).  
FT ACT\_SITE 205 205 Charge relay system (By similarity).  
FT DISULFID 36 137 By similarity.  
FT DISULFID 55 71 By similarity.  
FT DISULFID ? 239 By similarity.  
FT DISULFID 144 211 By similarity.  
FT DISULFID 176 190 By similarity.  
FT DISULFID 201 226 By similarity.  
FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;  
Query Match 100.0%; Score 40; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No; 6.4; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;  
Qy 1 SLLPLQLIL 9  
Db 4 SLLPLQLIL 12  
RESULT 3  
Q8NSN9 ID Q8NSN9 PRELIMINARY; PRT; 253 AA.  
AC Q8NSN9;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Stratum corneum chymotryptic enzyme, preproprotein.  
GN Name=KLK7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";   
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -I- SIMILARITY: Belongs to peptidase family S1.  
CC EMBL; BC032005; AAH32005.1; -;  
CC HSSP; P00760; IE2X.  
CC GO; GO:0004263; F:chymotrypsin activity; IEA.  
CC GO; GO:0008233; F:peptidase activity; IEA.  
CC GO; GO:0004295; F:trypsin activity; IEA.  
CC GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.

```

DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1;
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 27608 MW; 2D58B6A41B22A668 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9
DB 4 SLLPLQL 12

RESULT 4
Q9V2D5 PRELIMINARY; PRT; 146 AA.
AC Q9V2D5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PAB0088;
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Poch O., Priour D., Querellon J., Ripp R., Thierry J.-C.,
RA van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
archaeon Pyrococcus abyssi."
RL Mol. Microbiol. 47:1495-1512(2003).
DR EMBL; AJ248283; CAB49063.1; --
DR PIR; H75201; H75201.
DR InterPro; IPR008537; DUF819.
DR Pfam; PF05684; DUF819; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 146 AA; 16092 MW; 7182941371258C1F CRC64;

Query Match 95.0%; Score 38; DB 2; Length 146;
Best Local Similarity 88.9%; Pred. No. 9.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9
DB 66 SLLPLQL 74

RESULT 5
Q21527 PRELIMINARY; PRT; 208 AA.
AC Q21527;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
GN Name=ND4;
OS Clethrionomys gapperi (Southern red-backed vole).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
OC Clethrionomys.
OX NCBI_TaxID=56223;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;

```

RA Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC021345; AAH21345.1; -  
 DR MGD; MGI:1913370; Fkbp11.  
 DR GO; GO:0005635; C:extracellular space; TAS.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 SQ SEQUENCE 73 AA; 7819 MW; 938F53399BF3C11 CRC64;

Query Match 85.0%; Score 34; DB 2; Length 73;  
 Best Local Similarity 87.5%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQL 9  
 |||||:  
 Db 6 LLLPLQLL 13

RESULT 7  
 Q8R5D6 PRELIMINARY; PRT; 104 AA.

AC Q8R5D6;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Fkbp11 protein.  
 GN Name=Fkbp11;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CZECH II; TISSUE=Mammary tumor;  
 EX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC022900; AAH22900.1; -  
 DR GO; GO:0005635; C:extracellular space; TAS.  
 DR InterPro; IPR001179; FkBP\_PPIase.  
 DR PROSITE; PS50059; FkBP\_PPIase; 1.  
 SQ SEQUENCE 104 AA; 11085 MW; 0534D57467566914 CRC64;

Query Match 85.0%; Score 34; DB 2; Length 104;  
 Best Local Similarity 87.5%; Pred. No. 46;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQL 9  
 |||||:  
 Db 6 LLLPLQLL 13

RESULT 9  
 FKIL MOUSE  
 ID FKIL MOUSE STANDARD; PRT; 201 AA.

AC Q9D1M7; Q9CRE4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FK506 binding protein 11 precursor (EC 5.2.1.8) (Peptidyl-prolyl cis-  
 DE trans isomerase) (PPIase) (Rotamase) (19 kDa FK506-binding protein)  
 DE (FKBP-19).

Qy 2 LLLPLQL 9  
 |||||:  
 Db 6 LLLPLQLL 13

RESULT 8  
 Q6PKE2 PRELIMINARY; PRT; 138 AA.

AC Q6PKE2;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Fkbp11 protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;  
 EX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;  
 RA Strausberg R.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC002311; AAH02311.1; -  
 DR GO; GO:0005635; C:extracellular space; TAS.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR InterPro; IPR001179; FkBP\_PPIase.  
 DR PROSITE; PS50059; FkBP\_PPIase; 1.  
 SQ SEQUENCE 138 AA; 15105 MW; C138B8B0BFDF59D CRC64;

Query Match 85.0%; Score 34; DB 2; Length 138;  
 Best Local Similarity 87.5%; Pred. No. 61;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQL 9  
 |||||:  
 Db 6 LLLPLQLL 13

GN OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RP [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=22354683; PubMed=12468851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Oshio N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Watanston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smalus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: PPIases accelerate the folding of proteins during  
 CC protein synthesis.  
 CC -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline  
 CC (omega=0).  
 CC -!- SIMILARITY: Belongs to the FKBP-type PPIase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AF033340; AAC63945.1; -

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AK003331; BAB22719.1; -  
 CC EMBL; AK019132; BAB31559.1; -  
 CC EMBL; BC037596; AAH37596.1; -  
 CC HSSP; P20071; 1TCO.  
 CC MGD; MGI:1913370; Fkbp11.  
 CC InterPro; IPR001179; FKBP\_PPIase.  
 CC Pfam; PF00254; FKBP\_C; 1.  
 CC PROSITE; PS00509; FKBP\_PPIASE; 1.  
 CC Isomerase; Rotamase; Signal.  
 CC SIGNAL 1 27 Potential.  
 CC CHAIN 28 201 FK506 binding protein 11.  
 CC DOMAIN 57 144 PPIase, FKBP-type.  
 CC CONFLICT 53 53 S -> F (in Ref. 1; BAB31559).  
 CC CONFLICT 198 198 S -> R (in Ref. 1; BAB31559).  
 CC SEQUENCE 201 AA; 22137 MW; 94D955C57264BD82 CRC64;  
 Query Match 85.0%; Score 34; DB 1; Length 201;  
 Best Local Similarity 87.5%; Pred. No. 89;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 LLLPLQLL 9  
 Db 6 LLLPLQLL 13  
 ID IHA TRIVU STANDARD; PRT; 361 AA.  
 AC 077755;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DE 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Inhibin alpha chain precursor.  
 GN Name=INHA;  
 OS Trichosurus vulpecula (Brush-tailed possum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.  
 OX NCBI\_TaxID=9337;  
 RP [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=99027340; PubMed=9801457;  
 RA Vamontfort D., Fidler A.E., Heath D.A., Lawrence S.B., Tisdall D.J.,  
 RA Greenwood P.J., McNatty K.;  
 RT "cDNA sequence analysis, gene expression and protein localisation of  
 RT the inhibin alpha subunit of Australian brushtail possum (Trichosurus  
 RT vulpecula)";  
 RL J. Mol. Endocrinol. 21:141-152 (1998).  
 CC -!- FUNCTION: Inhibins and activins inhibit and activate,  
 CC respectively, the secretion of folliculin by the pituitary gland.  
 CC Inhibins/activins are involved in regulating a number of diverse  
 CC functions such as hypothalamic and pituitary hormone secretion,  
 CC gonadal hormone secretion, germ cell development and maturation,  
 CC erythroid differentiation, insulin secretion, nerve cell survival,  
 CC embryonic axial development or bone growth, depending on their  
 CC subunit composition. Inhibins appear to oppose the functions of  
 CC activins.  
 CC -!- SUBUNIT: Dimeric, linked by one or more disulfide bonds. Inhibin A  
 CC is a dimer of alpha and beta-A. Inhibin B is a dimer of alpha and  
 CC beta-B.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AF033340; AAC63945.1; -

DR GO: GO:0005576; C:extracellular; ISS.  
 DR GO: GO:0017106; F:activin inhibitor activity; ISS.  
 DR GO: GO:0005125; F:cytokine activity; ISS.  
 DR GO: GO:0008083; F:growth factor activity; ISS.  
 DR GO: GO:0005179; F:hormone activity; ISS.  
 DR GO: GO:0005515; F:protein binding; ISS.  
 DR GO: GO:0007050; P:cell cycle arrest; ISS.  
 DR GO: GO:0030154; P:cell differentiation; ISS.  
 DR GO: GO:0007165; P:cell growth and/or maintenance; ISS.  
 DR GO: GO:0007267; P:cell-cell signaling; ISS.  
 DR GO: GO:0030218; P:erythrocyte differentiation; ISS.  
 DR GO: GO:0042541; P:hemoglobin biosynthesis; ISS.  
 DR GO: GO:0006917; P:induction of apoptosis; ISS.  
 DR GO: GO:0045578; P:negative regulation of B-cell differentiation; ISS.  
 DR GO: GO:0045785; P:negative regulation of cell cycle; ISS.  
 DR GO: GO:0046882; P:negative regulation of follicle-stimulating. . .; ISS.  
 DR GO: GO:0045077; P:negative regulation of interferon-gamma bio. . .; ISS.  
 DR GO: GO:0045650; P:negative regulation of macrophage different. . .; ISS.  
 DR GO: GO:0042326; P:negative regulation of phosphorylation; ISS.  
 DR GO: GO:0007399; P:neurogenesis; ISS.  
 DR GO: GO:0001541; P:ovarian follicle development; ISS.  
 DR GO: GO:0046881; P:positive regulation of follicle-stimulating. . .; ISS.  
 DR InterPro: IPR002405; Inhibin\_alpha.  
 DR InterPro: IPR001839; TGFb.  
 DR Pfam: PF00019; TGF\_beta; 1.  
 DR PRINTS: PR00669; INHIBINA.  
 DR ProDom: PD000357; TGFb; 1.  
 DR SMART: SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Glycoprotein; Growth factor; Hormone; Signal.  
 FT SIGNAL 1 21  
 FT PROPEP 22 230  
 FT CHAIN 231 361  
 FT DISULFID 260 323  
 FT DISULFID 289 358  
 FT DISULFID 293 360  
 FT DISULFID 322 322  
 FT CARBOHYD 48 48  
 FT CARBOHYD 144 144  
 FT CARBOHYD 266 266  
 SQ SEQUENCE 361 AA; 38945 MW; D661CDF93CDA87D CRC64;  
 Query Match 85.0%; Score 34; DB 1; Length 361;  
 Best Local Similarity 87.5%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LLLPLQL 9  
 Db 4 LLLPLQLL 11  
 RESULT 11  
 Q749Q7 PRELIMINARY; PRT; 373 AA.  
 AC Q749Q7  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE Membrane protein, putative.  
 GN OrderedLocusNames=GSU2685;  
 OS Geobacter sulfurreducens.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;  
 OC Geobacteraceae; Geobacter.  
 OX NCBI\_TaxID=35554;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCA / ATCC 51573;  
 RX PubMed=14671304; DOI=10.1126/science.1089727;  
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C., Heidelberg J.F., Wu D., Wu M., Ward N.L., Seaman M.J., Dodson R.J., Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S., Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,

RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A., Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R., Van Aken S.E., Lovley D.R., Fraser C.M.;  
 "Genome of Geobacter sulfurreducens: metal reduction in subsurface environments.";  
 RL Science 302:1967-1969(2003).  
 DR EMBL; AE017180; AAR36057.1; -.  
 DR TIGR; GSU2685; -.  
 DR InterPro: IPR000412; ABC\_2.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 373 AA; 40412 MW; 94B9D6APF570D4A2 CRC64;  
 Query Match 85.0%; Score 34; DB 2; Length 373;  
 Best Local Similarity 87.5%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LLLPLQL 9  
 Db 293 LLLPLQL 300  
 RESULT 12  
 Q6N500 PRELIMINARY; PRT; 148 AA.  
 AC Q6N500;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE Hypothetical protein precursor.  
 GN OrderedLocusNames=RPA3183;  
 OS Rhodopseudomonas palustris.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Rhodopseudomonas.  
 OX NCBI\_TaxID=1076;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CGA009 / ATCC BAA-98;  
 RX PubMed=14704707; DOI=10.1038/nbt923;  
 RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L., Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C., Harrison F.H., Gibson J., Harwood C.S.;  
 "Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris.";  
 RL Nat. Biotechnol. 22:55-61(2004).  
 DR EMBL; BX572603; CAE28624.1; -.  
 DR InterPro: IPR000975; Interleukin\_1.  
 DR PROSITE; PS00253; INTERLEUKIN\_1; UNKNOWN 1.  
 KW Complete proteome; Hypothetical protein; Signal.  
 FT SIGNAL 1 22  
 SQ SEQUENCE 148 AA; 16860 MW; 9501B7C6C2808F45 CRC64;  
 Query Match 82.5%; Score 33; DB 2; Length 148;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SLLPLQL 9  
 Db 7 SLLPLMLL 15  
 RESULT 13  
 Q6GNT6 PRELIMINARY; PRT; 302 AA.  
 AC Q6GNT6;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE MGC80882 protein.  
 GN Name=MGC80882;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McWeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family
DR EMBL; BC073416; AAH73416.1; --
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH short.
DR InterPro; IPR002347; Adh short_C2.
DR Pfam; PF00106; adh short; 1.
DR PRINTS; PR00081; GDRRDH.
DR PRINTS; PR00080; SDRPAMILY.
KW Oxidoreductase.
SQ SEQUENCE 302 AA; 33350 MW; 6F9503C68340313A CRC64;

Query Match 82.5%; Score 33; DB 2; Length 302;
Best Local Similarity 87.5%; Pred. No. 2.1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLPLQIL 9
DB 10 LLLPLQIL 17
|||||

RESULT 14
YQ47 XANAC STANDARD; PRT; 339 AA.
AC Q8PLX8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical UPF0190 protein XAC1647.
GN OrderedLocNames=XAC1647;
OS Xanthomonas axonopodis (pv. citri).

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
NCBI_TaxID=8355;
SEQUENCE FROM N.A.
TISSUE=Embryo;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McWeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
SEQUENCE FROM N.A.
TISSUE=Embryo;
MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
"Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
Dev. Dyn. 225:384-391(2002).
SEQUENCE FROM N.A.
TISSUE=Embryo;
MEDLINE=22024217; DOI=10.1038/417459a;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
Kasuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezsa R.I.D.,
Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
Setubal J.C., Kitajima J.P.;
"Comparison of the genomes of two Xanthomonas pathogens with differing
host specificities.";
Nature 417:459-463(2002).
-!- SIMILARITY: Belongs to the UPF0190 family.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; AE011796; AAM36515.1; --
DR HAMAP; MF_01206; 1
DR InterPro; IPR000572; Oxidored_molyb.
DR InterPro; IPR006311; Tat.
DR Pfam; PF00174; Oxidored_molyb; 1.
DR TIGRFAMs; TIGR01409; TAT signal_seq; 1.
KW Complete proteome; Hypothetical_protein.
SQ SEQUENCE 339 AA; 37901 MW; B926DFC9AA849438 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 339;
Best Local Similarity 88.9%; Pred. No. 2.4e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLLPLQL 9
DB 5 SLLPLQL 13
|||||

RESULT 15
Q8VHY9 PRELIMINARY; PRT; 370 AA.
AC Q8VHY9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DAUNORUBICIN RESISTANCE TRANSMEMBRANE PROTEIN.
GN OrderedLocNames=BMEI0656;
OS Brucella melitensis.
OC Brucella; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,

```

RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Hasselkorn R., Kyripides N.C., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).  
 DR EMBL; AB009507; AAL51837.1; -.  
 DR PIR; AB3334; AB3334.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 370 AA; 40369 MW; 819AAD684305CB60 CRC64;

Query Match 82.5%; Score 33; DB 2; Length 370;  
 Best Local Similarity 75.0%; Pred. NO. 2.6e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLPLQL 9  
 :|||||:  
 Db 293 ILLPLQVL 300

Search completed: July 13, 2005, 17:29:25  
 Job time : 63.2222 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-36

Perfect score: 40

Sequence: 1 SLLPLQLL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	3	US-09-502-600-36
2	40	100.0	9	4	US-09-918-243-36
3	40	100.0	253	2	US-08-557-146-2
4	40	100.0	253	2	US-08-824-874-3
5	40	100.0	253	2	US-09-154-344-2
6	40	100.0	253	3	US-08-930-188-2
7	40	100.0	253	3	US-09-210-084-3
8	40	100.0	253	4	US-09-764-762-3
9	40	100.0	253	5	PCT-US96-04294-2
10	40	100.0	265	4	US-09-949-016-7716
11	36	90.0	9	3	US-09-502-600-33
12	36	90.0	9	4	US-09-918-243-33
13	34	85.0	812	4	US-09-489-039A-12075
14	34	80.0	9	3	US-09-502-600-35
15	32	80.0	9	3	US-09-502-600-116
16	32	80.0	9	4	US-09-918-243-35
17	32	80.0	9	4	US-09-918-243-116
18	32	80.0	295	4	US-09-270-767-39889
19	32	80.0	295	4	US-09-270-767-55106
20	32	80.0	557	4	US-09-489-039A-10804
21	31	77.5	181	4	US-09-543-681A-5937
22	31	77.5	307	3	US-09-177-249-13
23	31	77.5	307	4	US-09-812-283-13
24	31	77.5	472	4	US-09-543-681A-5049
25	31	77.5	1279	4	US-09-170-496D-293
26	31	77.5	1279	4	US-09-364-425B-58
27	30	75.0	135	2	US-08-383-621-5

28	30	75.0	135	3	US-08-459-906-5	Sequence 5, Appli
29	30	75.0	169	4	US-09-328-352-7861	Sequence 7861, Ap
30	30	75.0	231	4	US-09-252-991A-22663	Sequence 22663, A
31	30	75.0	308	4	US-09-252-991A-27921	Sequence 27921, A
32	30	75.0	333	4	US-09-949-016-6128	Sequence 6128, Ap
33	30	75.0	340	4	US-09-949-016-10615	Sequence 10615, A
34	30	75.0	404	4	US-09-489-039A-8947	Sequence 8947, Ap
35	30	75.0	436	4	US-09-489-039A-11785	Sequence 11785, A
36	30	75.0	463	2	US-08-677-049-6	Sequence 6, Appli
37	30	75.0	474	4	US-09-489-039A-10791	Sequence 10791, A
38	30	75.0	542	4	US-09-489-847-323	Sequence 323, Ap
39	30	75.0	549	2	US-08-676-279-57	Sequence 57, Appl
40	30	75.0	561	4	US-09-248-796A-25215	Sequence 25215, A
41	30	75.0	599	4	US-09-949-016-9866	Sequence 9866, Ap
42	30	75.0	1098	1	US-07-777-715-7	Sequence 7, Appli
43	30	75.0	1098	1	US-08-170-126-2	Sequence 2, Appli
44	30	75.0	1098	3	US-08-954-418-2	Sequence 2, Appli
45	29	72.5	21	3	US-08-753-007A-24	Sequence 24, Appl

## ALIGNMENTS

RESULT 1  
US-09-502-600-36  
; Sequence 36, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 36  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 4-12 of the SCCE protein  
US-09-502-600-36

Query Match 100.0%; Score 40; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLLPLQLL 9  
Db 1 SLLPLQLL 9

RESULT 2  
US-09-918-243-36  
; Sequence 36, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 36  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 4-12 of the SCCE protein
US-09-918-243-36

Query Match      100.0%; Score 40; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9
    |||||
Db 1 SLLPLQL 9

RESULT 3
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hanson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 14-DEC-1995
; APPLICATION NUMBER: US/08/557,146
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2

Query Match      100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9
    |||||
Db 4 SLLPLQL 12

RESULT 4
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 16-SEP-1998
; APPLICATION NUMBER: US/09/154,344
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; 
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-08-824-874-3

Query Match      100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLLPLQL 9
    |||||
Db 4 SLLPLQL 12

RESULT 5
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hanson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; 
```

APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-154-344-2

Query Match 100.0%; Score 40; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.6; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12

RESULT 6  
US-08-930-188-2  
Sequence 2, Application US/08930188  
Patent No. 6093397  
GENERAL INFORMATION:  
APPLICANT: Dixon, Eric P.  
ADDRESSEE: Johnstone, Edward M.  
APPLICANT: Little, Sheila P.  
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
RELATED NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930,188  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/416,257  
FILING DATE: 04-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Blalock, Donna K.  
REGISTRATION NUMBER: 38,082  
REFERENCE/DOCKET NUMBER: X9239  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-277-1090  
TELEFAX: 317-276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-930-188-2

Query Match 100.0%; Score 40; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.6; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12  
RESULT 7  
US-09-210-084-3  
Sequence 3, Application US/09210084  
Patent No. 6197511  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,084  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,874  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0252 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 532504  
US-09-210-084-3

Query Match 100.0%; Score 40; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.6; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12

RESULT 8  
US-09-764-762-3  
Sequence 3, Application US/09764762  
Patent No. 6472195  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

      Query Match      100.0%; Score 40; DB 4; Length 253;
      Best Local Similarity 100.0%; Pred. No. 2.6;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9
Db 4 SLLPLQL 12

RESULT 9
PCT-US96-04294-2
;
; GENERAL INFORMATION:
; APPLICANT: Dixon, Eric P.
; APPLICANT: Johnstone, Edward M.
; APPLICANT: Little, Sheila P.
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
; TITLE OF INVENTION: RELATED NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/04294
; FILING DATE:
; CLASSIFICATION:
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/416,257
; FILING DATE: 04-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Blalock, Donna K.
; REGISTRATION NUMBER: 38,082
; REFERENCE/DOCKET NUMBER: X9239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-277-1090
; TELEFAX: 317-276-3861
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-04294-2

      Query Match      100.0%; Score 40; DB 5; Length 253;
      Best Local Similarity 100.0%; Pred. No. 2.6;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9
Db 4 SLLPLQL 12

RESULT 10
US-09-949-016-7716
;
; Sequence 7716, Application US/09949016
; Patent No. 6812339
;
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7716
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Human
;
; US-09-949-016-7716

      Query Match      100.0%; Score 40; DB 4; Length 265;
      Best Local Similarity 100.0%; Pred. No. 2.7;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9
Db 16 SLLPLQL 24

RESULT 11
US-09-502-600-33
;
; Sequence 33, Application US/09502600A
; Patent No. 6294344
;
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
```

```
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-502-600-33

Query Match
Best Local Similarity 90.0%; Score 36; DB 3; Length 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQIL 9
Db 1 LLLPLQIL 8

RESULT 12
US-09-918-243-33
; Sequence 33, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-918-243-33

Query Match
Best Local Similarity 90.0%; Score 36; DB 4; Length 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQIL 9
Db 1 LLLPLQIL 8

RESULT 13
US-09-489-039A-12075
; Sequence 12075, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12075
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12075
```

```
Query Match
Best Local Similarity 85.0%; Score 34; DB 4; Length 812;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQIL 9
Db 429 LLLPLQLL 436

RESULT 14
US-09-502-600-35
; Sequence 35, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-502-600-35

Query Match
Best Local Similarity 80.0%; Score 32; DB 3; Length 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LLLPLQIL 9
Db 1 LLLPLQIL 7

RESULT 15
US-09-502-600-116
; Sequence 116, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 2-10 of the SCCE protein
US-09-502-600-116

Query Match
Best Local Similarity 80.0%; Score 32; DB 3; Length 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQ 7
Db 3 SLLPLQ 9

Search completed: July 13, 2005, 17:34:28
Job time : 20.4444 secs
```

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds  
(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-36

Perfect score: 40

Sequence: 1 SLLPLQL 9

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 366332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pcp.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	9 US-09-918-243-36	Sequence 36, Appl
2	40	100.0	9	9 US-09-905-083-36	Sequence 36, Appl
3	40	100.0	9	15 US-10-372-521-36	Sequence 36, Appl
4	40	100.0	9	16 US-10-831-075-36	Sequence 36, Appl
5	40	100.0	198	15 US-10-262-511-96	Sequence 92, Appl
6	40	100.0	250	15 US-10-262-511-92	Sequence 92, Appl
7	40	100.0	253	9 US-09-888-615-98	Sequence 98, Appl
8	40	100.0	253	9 US-09-764-762-3	Sequence 3, Appl
9	40	100.0	253	14 US-10-071-214-2	Sequence 2, Appl
10	40	100.0	253	14 US-10-071-214-48	Sequence 48, Appl
11	40	100.0	253	14 US-10-264-283-90	Sequence 90, Appl

12	40	100.0	253	15 US-10-295-027-498	Sequence 498, Appl
13	40	100.0	253	15 US-10-173-999-48	Sequence 48, Appl
14	40	100.0	253	16 US-10-408-765A-639	Sequence 639, Appl
15	40	100.0	253	16 US-10-643-795A-95	Sequence 95, Appl
16	40	100.0	253	17 US-10-948-518-95	Sequence 95, Appl
17	40	100.0	253	17 US-10-868-490A-1	Sequence 1, Appl
18	40	100.0	257	15 US-10-344-394-38	Sequence 38, Appl
19	36	90.0	9	9 US-09-918-243-33	Sequence 33, Appl
20	36	90.0	9	9 US-09-905-083-33	Sequence 33, Appl
21	36	90.0	9	15 US-10-372-521-33	Sequence 33, Appl
22	36	90.0	9	16 US-10-831-075-33	Sequence 33, Appl
23	35	87.5	138	15 US-10-425-114-53148	Sequence 53148, A
24	35	87.5	868	15 US-10-369-493-10255	Sequence 10255, A
25	34	85.0	56	15 US-10-424-599-164977	Sequence 164977, A
26	34	85.0	70	15 US-10-424-599-207950	Sequence 207950, A
27	34	85.0	379	16 US-10-437-963-111334	Sequence 111334, A
28	33	82.5	46	15 US-10-424-599-265948	Sequence 265948, A
29	33	82.5	86	15 US-10-424-599-199921	Sequence 199921, A
30	33	82.5	137	16 US-10-425-115-264649	Sequence 264649, A
31	33	82.5	394	15 US-10-282-122A-45727	Sequence 45727, A
32	33	82.5	397	16 US-10-408-765A-2842	Sequence 2842, Appl
33	33	82.5	804	14 US-10-236-055A-28	Sequence 28, Appl
34	32	80.0	9	9 US-09-918-243-35	Sequence 35, Appl
35	32	80.0	9	9 US-09-918-243-116	Sequence 116, Appl
36	32	80.0	9	9 US-09-905-083-35	Sequence 35, Appl
37	32	80.0	9	9 US-09-905-083-116	Sequence 116, Appl
38	32	80.0	9	15 US-10-372-521-35	Sequence 35, Appl
39	32	80.0	9	15 US-10-372-521-116	Sequence 116, Appl
40	32	80.0	9	16 US-10-831-075-35	Sequence 35, Appl
41	32	80.0	9	16 US-10-831-075-116	Sequence 116, Appl
42	32	80.0	54	16 US-10-425-115-368614	Sequence 368614, A
43	32	80.0	74	15 US-10-424-599-252879	Sequence 252879, A
44	32	80.0	78	17 US-10-660-811A-173	Sequence 173, Appl
45	32	80.0	80	15 US-10-424-599-167498	Sequence 167498, A

## ALIGNMENTS

### RESULT 1

US-09-918-243-36  
; Sequence 36, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 36  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 4-12 of the SCCE protein  
US-09-918-243-36

Query Match 100.0%; Score 40; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9

Db 1 SLLPLQL 9

### RESULT 2

US-09-905-083-36  
 ; Sequence 36, Application US/09905083  
 ; Patent No. US20020146708A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, Timothy J.  
 ; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
 ; TITLE OF INVENTION: Ovarian Cancer  
 ; FILE REFERENCE: D6223CIP/C/Div  
 ; CURRENT APPLICATION NUMBER: US/09/905,083  
 ; CURRENT FILING DATE: 2001-07-13  
 ; PRIOR APPLICATION NUMBER: US 09/502,600  
 ; PRIOR FILING DATE: 2000-02-11  
 ; NUMBER OF SEQ ID NOS: 136  
 ; SEQ ID NO 36  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CHAIN  
 ; OTHER INFORMATION: Residues 4-12 of the SCCE protein  
 US-09-905-083-36

Query Match 100.0%; Score 40; DB 9; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
 Db 1 SLLPLQL 9

RESULT 3  
 US-10-372-521-36  
 ; Sequence 36, Application US/10372521  
 ; Publication No. US20030223973A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, Timothy J.  
 ; APPLICANT: Cannon, Martin J.  
 ; APPLICANT: Santin, Alessandro  
 ; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
 ; FILE REFERENCE: D6223CIP/C/D/CIP2  
 ; CURRENT APPLICATION NUMBER: US/10/372,521  
 ; CURRENT FILING DATE: 2003-02-21  
 ; PRIOR APPLICATION NUMBER: US 09/918,243  
 ; PRIOR FILING DATE: 2001-07-30  
 ; NUMBER OF SEQ ID NOS: 136  
 ; SEQ ID NO 36  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CHAIN  
 ; OTHER INFORMATION: Residues 4-12 of the SCCE protein  
 US-10-372-521-36

Query Match 100.0%; Score 40; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
 Db 1 SLLPLQL 9

RESULT 4  
 US-10-831-075-36  
 ; Sequence 36, Application US/10831075  
 ; Publication No. US20040224891A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, Timothy J.  
 ; APPLICANT: Cannon, Martin J.  
 ; APPLICANT: Santin, Alessandro  
 ; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP3  
 ; CURRENT APPLICATION NUMBER: US/10/831,075  
 ; CURRENT FILING DATE: 2004-04-23  
 ; PRIOR APPLICATION NUMBER: US 10/372,521  
 ; PRIOR FILING DATE: 2003-02-21  
 ; NUMBER OF SEQ ID NOS: 140  
 ; SEQ ID NO 36  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CHAIN  
 ; OTHER INFORMATION: Residues 4-12 of the SCCE protein  
 US-10-831-075-36

Query Match 100.0%; Score 40; DB 16; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
 Db 1 SLLPLQL 9

RESULT 5  
 US-10-262-511-96  
 ; Sequence 96, Application US/10262511  
 ; Publication No. US20040038223A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Ju, Jingfang  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Guo, Xiaojia (Sasha)  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Ort, Tatiana  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Zerhusen, Bryan D.  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Pena, Carol E. A.  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Shimkets, Richard A.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Berghs, Constance  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-462C  
 ; CURRENT APPLICATION NUMBER: US/10/262,511  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/373,815  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,642  
 ; PRIOR FILING DATE: 2002-05-17  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 96  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-96

Query Match 100.0%; Score 40; DB 15; Length 198;  
Best Local Similarity 100.0%; Pred. No. 20; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12

## RESULT 6

US-10-262-511-92  
; Sequence 92, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Rotherberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; PRIOR FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 92  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-92

Query Match 100.0%; Score 40; DB 15; Length 250;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
Db 1 SLLPLQL 9

## RESULT 7

US-09-888-615-98  
; Sequence 98, Application US/09888615  
; Patent No. US20020064856A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: CAENEPEEL, SEAN  
; APPLICANT: CHARYDCZAK, GLEN  
; APPLICANT: MANNING, GERARD  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: NOVEL PROTEASES  
; FILE REFERENCE: 038602/1214  
; CURRENT APPLICATION NUMBER: US/09/888,615  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,047  
; PRIOR FILING DATE: 2000-06-26  
; NUMBER OF SEQ ID NOS: 150  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 98  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-888-615-98

Query Match 100.0%; Score 40; DB 9; Length 253;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
Db 4 SLLPLQL 12

## RESULT 8

US-09-764-762-3  
; Sequence 3, Application US/09764762  
; Patent No. US20020068341A1  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: 09/210,084  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 532504  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-764-762-3  
Query Match 100.0%; Score 40; DB 9; Length 253;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
| | | | |  
Db 4 SLLPLQL 12

RESULT 9  
US-10-071-214-2  
; Sequence 2, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGELRUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-071-214-2

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
| | | | |  
Db 4 SLLPLQL 12

RESULT 10  
US-10-071-214-48  
; Sequence 48, Application US/10071214  
; Publication No. US20030066099A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSSON, Lennart  
; APPLICANT: EGELRUD, Torbjorn  
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN  
; FILE REFERENCE: HANSSON=3A  
; CURRENT APPLICATION NUMBER: US/10/071,214  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: US 60/267,422  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: DK PA 2001 00218  
; PRIOR FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 48  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Deduced amino acid sequence from the C-terminal part of SCCE from  
US-10-071-214-48

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
| | | | |  
Db 4 SLLPLQL 12

RESULT 11  
US-10-264-283-90  
; Sequence 90, Application US/10264283  
; Publication No. US20030144494A1  
; GENERAL INFORMATION:  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.590  
; CURRENT APPLICATION NUMBER: US/10/264,283  
; CURRENT FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 90  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-283-90

Query Match 100.0%; Score 40; DB 14; Length 253;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9  
| | | | |  
Db 4 SLLPLQL 12

RESULT 12  
US-10-295-027-498  
; Sequence 498, Application US/10295027

```
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9
Db 4 SLLPLQL 12

RESULT 13
US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
```

```
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-48

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9
Db 4 SLLPLQL 12

RESULT 14
US-10-408-765A-639
; Sequence 639, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary W.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088,465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 639
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-639

Query Match      100.0%; Score 40; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLLPLQL 9
Db 4 SLLPLQL 12

RESULT 15
US-10-643-795A-95
; Sequence 95, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN PRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026F1-US
; CURRENT APPLICATION NUMBER: US/10/643,795A
```

```
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 95
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-643-795A-95
```

```
Query Match      100.0%; Score 40; DB 16; Length 253;
Best Local Similarity 100.0%; Pred.No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 SLLPLQL 9
         |||||
Db      4 SLLPLQL 12
```

Search completed: July 13, 2005, 18:23:40  
Job time : 65.8889 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083A-80

Perfect score: 50

Sequence: 1 GPLVCRGTL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq 16Dec04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	9	4	Aae08285 Human str
2	50	100.0	9	8	Adr68841 Human str
3	50	100.0	181	6	Ada05738 Human NOV
4	50	100.0	181	8	Adh62902 Human NOV
5	50	100.0	198	6	Ada05736 Human NOV
6	50	100.0	198	8	Adh62900 Human NOV
7	50	100.0	224	6	Ada05744 Human NOV
8	50	100.0	224	8	Adh62908 Human NOV
9	50	100.0	225	4	Aab98502 Human str
10	50	100.0	247	6	Ada05742 Human str
11	50	100.0	247	8	Adh62906 Human NOV
12	50	100.0	250	6	Ada05732 Human NOV
13	50	100.0	250	8	Adh62896 Human NOV
14	50	100.0	252	6	Ada05734 Human NOV
15	50	100.0	252	8	Adh62898 Human NOV
16	50	100.0	253	2	Aar67888 Human str
17	50	100.0	253	2	Aaw05383 Human amy
18	50	100.0	253	5	Abb84421 Human SCC
19	50	100.0	253	5	Abb84406 Human SCC
20	50	100.0	253	5	Aau82740 Amino aci
21	50	100.0	253	6	Abu07440 Protein d
22	50	100.0	253	6	Abu07471 Protein d
23	50	100.0	253	6	Abt58471 Human str
24	50	100.0	253	7	Adh80484 Ovarian c
25	50	100.0	253	7	Adj68833 Human hea

26	50	100.0	253	7	Adn39180 Cancer/an
27	50	100.0	253	8	Adl06515 Human tum
28	50	100.0	253	8	Adn04182 Antipsoi
29	50	100.0	253	8	Adr72880 Human ova
30	50	100.0	257	3	Abb21326 Human HSC
31	47	94.0	249	5	Abb84420 Porcine S
32	45	90.0	247	4	Aau86677 Novel hum
33	45	90.0	247	4	Aau23217 Novel hum
34	45	90.0	247	4	Aau23752 Novel hum
35	45	90.0	247	4	Aau17043 Human nov
36	45	90.0	247	7	Adb60011 Connectiv
37	45	90.0	250	3	Abb21298 Human CLK
38	45	90.0	250	5	Abp64969 Human pro
39	45	90.0	250	6	Abt55400 Amino aci
40	45	90.0	250	8	Adn05516 Antipsoi
41	45	90.0	251	4	Aau16971 Human nov
42	45	90.0	259	5	Adi16686 Human NOV
43	45	90.0	259	8	Adn42340 Human nov
44	45	90.0	296	3	Abb21297 Human CLK
45	44	88.0	165	8	Abm84664 Human dia

## ALIGNMENTS

### RESULT 1

ID	AAE08285	standard; peptide; 9 AA.
XX	AC	AAE08285;
XX	DT	01-NOV-2001 (first entry)
XX	DE	Human stratum corneum chymotrypsin enzyme peptide #50 (residues 207-215).
XX	KW	Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
XX	KW	cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
XX	OS	antisense therapy; malignant hyperplasia.
XX	PN	Homo sapiens.
XX	PD	WO200159158-A1.
XX	PF	16-AUG-2001.
XX	PR	07-FEB-2001; 2001WO-US003977.
XX	PI	11-FEB-2000; 2000US-00502600.
XX	PA	(UYAR-) UNIV ARKANSAS.
XX	PI	O'brien TJ;
XX	DR	WPI; 2001-514676/56.
XX	PT	Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.
XX	PS	Claim 25; Page 113; 127pp; English.

The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

Sequence 9 AA;

Query Match 100.0%; Score 50; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||||  
Db 1 GPLVCRGTL 9

## RESULT 2

AD68841  
ID AD68841 standard; peptide; 9 AA.

XX AC AD68841;  
XX

DT 02-DEC-2004 (first entry)

DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:80.

XX serine protease; stratum corneum chymotrytic enzyme; SCCE;  
KW immune response; ovarian cancer; lung cancer; prostate cancer;  
KW pancreatic cancer; colon cancer.

XX OS Homo sapiens.

XX PN WO2004075723-A2.

XX PD 10-SEP-2004.

XX PF 20-FEB-2004; 2004WO-US005134.

XX PR 21-FEB-2003; 2003US-00372521.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'brien TJ, Cannon MJ, Santin A;

XX DR WPI; 2004-653294/63.

XX Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for  
PT vaccinating an individual against SCCE, and in monitoring and developing  
PT immunotherapies for ovarian and other malignancies.

XX Claim 5; SEQ ID NO 80; 117pp; English.

XX The invention relates to a novel method for vaccinating an individual  
CC against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating  
CC an individual with a SCCE peptide, which elicits an immune response in  
CC the individual. A peptide of the invention acts as a stratum corneum  
CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
CC an individual against SCCE, particularly an individual having, suspected  
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
CC an individual, such as ovarian, breast, lung, colon, prostate, or  
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
CC peptides are also useful in the monitoring and development of  
CC immunotherapies for ovarian and other malignancies. The present sequence  
CC represents a peptide fragment of serine protease SCCE (stratum corneum  
CC chymotrytic enzyme).

XX Sequence 9 AA;

Query Match 100.0%; Score 50; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||||  
Db 1 GPLVCRGTL 9

## RESULT 3

ADA05738  
ID ADA05738 standard; protein; 181 AA.

XX AC ADA05738;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV18d protein SEQ ID NO:98.  
XX  
KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipase; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031373.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329141P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339286P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.  
XX (CURA-) CURAGEN CORP.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman I, Zerhusen BD, Anderson DM, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2003-381626/36.  
XX N-PSDB; ADA05737.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX Claim 1; Page 171; 586pp; English.  
XX  
XX The present invention describes NOVX proteins, where X can be 1 to 55



CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and the  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 181 AA;

Query Match 100.0%; Score 50; DB 6; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPLVCRGTL 9  
 |||||  
 Db 133 GPLVCRGTL 141

# RESULT 4

ID ADN62902 standard; protein; 181 AA.

XX AC ADN62902;

XX DT 01-JUL-2004 (first entry)

XX DE Human NOV18d.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

XX OS Homo sapiens.

XX PN US2004038223-A1.

XX DT 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328039P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383811P.  
 PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J. A.

PA (KEKU/) KEKUDA R.

PA (JUJ/) JU J.

PA (LILL/) LI L.

PA (GUOX/) GUO X.

PA (PATT/) PATTURAJAN M.

PA (SPYT/) SPYTEK K. A.

PA (EDIN/) EDINGER S. R.

PA (ELLE/) ELLERMAN K.

PA (MALY/) MALYANKAR U. M.

PA (ORTT/) ORT T.

PA (GORM/) GORMAN L.

PA (ZERH/) ZERHUSEN B. D.

PA (ANDE/) ANDERSON D. W.

PA (ZHON/) ZHONG M.

PA (CAIT/) CATTERTON E.

PA (JIW/) JI W.

PA (MILL/) MILLER C. E.

PA (RST/) RASTELLI L.

PA (STON/) STONE D. J.

PA (PENA/) PENNA C. E. A.

PA (SHEN/) SHENOY S. G.

PA (SHIM/) SHIMKETS R. A.

PA (ROTH/) ROTHENBERG M. E.

PA (LEAC/) LEACH M. D.

PA (AGEE/) AGE E. M. L.

PA (BERG/) BERGHS C.

PA (DIP/) DIPIPPO V. A.

PA (EISE/) EISEN A.

PA (GANG/) GANGOLLI E. A.

PA (RIEG/) RIEGER D. K.

PA (SPAD/) SPADERNA S. K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytke KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2004-213931/20.

DR N-PSDB; ADN62901.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,

PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 98; 395pp; English.

PS The invention relates to isolated NOVX polypeptides and polynucleotides.

XX NOVX polypeptides and polynucleotides are used to prevent, diagnose or

CC treat a medical condition in human related to the aberrant expression and

CC activity of NOVX polypeptides. For example, NOVX polypeptides and

CC polynucleotides may be used to treat disorders associated with decreased

CC expression or activity of NOVX by supplementing the patient our

CC production or to rectify mutations. Conversely, antisense NA molecules

CC may be administered to down regulate expression of NOVX polypeptides by

CC binding with the cells own genes and preventing their expression. NOVX

CC polynucleotides and complementary sequences may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of

CC similar sequences in samples, and so which patients may be in need of

CC restorative therapy. NOVX polypeptides may also be used as antigens in

CC the production of antibodies and in assays to identify modulators

CC (agonists and antagonists) of the expression and activity of NOVX. The

CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be

CC used to modulate NOVX polynucleotide expression and activity of NOVX

CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as

CC diagnostic agents for detecting the presence of NOVX in samples. NOVX

CC polypeptides and polynucleotides may be used in this way to prevent,

CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious

CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative

CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,

CC haematopoietic disorders, and the various dyslipidaemias, metabolic

CC disturbances associated with obesity, the metabolic syndrome X and

CC wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence

XX represents the amino acid sequence of a human NOVX protein.

SQ Sequence 181 AA;

Query Match 100.0%; Score 50; DB 8; Length 181;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9

DB 133 GPLVCRGTL 141

RESULT 5

ADA05736

ID ADA05736 standard; protein; 198 AA.

AC ADA05736;

DT 06-NOV-2003 (first entry)

DE Human NOV18c protein SEQ ID NO:96.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytostatic; nootropic; neuroprotective;

KW antiparkinsonian; antilipaeamic; Gene therapy; human disease;

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

OS

XX WO2003029424-A2.

PN

XX 10-APR-2003.

PD

XX 02-OCT-2002; 2002WO-US031373.

PF

XX 02-OCT-2001; 2001US-0326483P.

PR

XX 05-OCT-2001; 2001US-0327435P.

PR

XX 09-OCT-2001; 2001US-0327449P.

PR

XX 09-OCT-2001; 2001US-0327917P.

PR

XX 09-OCT-2001; 2001US-0328029P.

PR

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0373884P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CEA, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

DR N-PSDB; ADA05735.

XX

PT New NOVX polypeptides and nucleic acids, useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX

PS Claim 1; Page 170; 586pp; English.

CC The present invention describes NOVX proteins, where X can be 1 to 55

CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described

CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or

CC aberrant physiological interactions of the polypeptide; (11) a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating

CC the activity of the polypeptide described above; (13) methods of treating

CC or preventing a pathology associated with the above polypeptide in a

CC mammal; and (14) a method for producing the above polypeptide. NOVX

CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian

CC and antilipaeamic activities, and can be used in gene therapy. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic

CC disorders such as diabetes or obesity, infections, cachexia, cancer,

CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

SQ Sequence 198 AA;

Query Match 100.0%; Score 50; DB 6; Length 198;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9

Db 152 GPLVCRGTL 160

RESULT 6

ADN62900

ID ADN62900 standard; protein; 198 AA.

XX AC ADN62900;

DT 01-JUL-2004 (first entry)

XX DE Human NOV18c.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

XX KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;

XX KW Alzheimer's disease; Parkinson's disease; immune disorder;

XX KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;

XX KW wasting disorder.

XX OS Homo sapiens.

XX PN US2004038223-A1.

XX PD 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 09-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

XX PR 19-APR-2002; 2002US-0373826P.

XX PR 19-APR-2002; 2002US-0373884P.

XX PR 22-APR-2002; 2002US-0374977P.

XX PR 16-MAY-2002; 2002US-0381037P.

XX PR 16-MAY-2002; 2002US-0381038P.

XX PR 16-MAY-2002; 2002US-0381042P.

XX PR 17-MAY-2002; 2002US-0381642P.

XX PR 28-MAY-2002; 2002US-0383656P.

XX PR 29-MAY-2002; 2002US-0383831P.

XX PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA

PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENNA/) PENNA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEACH/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIPV/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malvankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2004-213931/20.

DR N-PSDB; ADN62899.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 96; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic

CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 198 AA;

Query Match 100.0%; Score 50; DB 8; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPLVCRGTL 9  
Db 152 GPLVCRGTL 160  
|||||

RESULT 7  
ADA05744  
ID ADA05744 standard; protein; 224 AA.

AC ADA05744;

DT 06-NOV-2003 (first entry)

DE Human NOV18g protein SEQ ID NO:104.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

PN 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 198 AA;

Query Match 100.0%; Score 50; DB 8; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPLVCRGTL 9  
Db 152 GPLVCRGTL 160  
|||||

RESULT 7  
ADA05744  
ID ADA05744 standard; protein; 224 AA.

AC ADA05744;

DT 06-NOV-2003 (first entry)

DE Human NOV18g protein SEQ ID NO:104.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

PN 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CBA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dippio VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI: 2003-381626/36.  
DR N-PSDB; ADA05743.  
XX

PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX

PS Claim 1; Page 172; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC binds to the polypeptide described above; (7) methods for immunospecifically  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
CC and antilipemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
XX present invention.

XX Sequence 224 AA;

Query Match 100.0%; Score 50; DB 6; Length 224;

Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPLVCRGTL 9

Db 193 GPLVCRGTL 201

|||||

RESULT 8

ADN62908

ID ADN62908 standard; protein; 224 AA.

XX ADN62908;

XX 01-JUL-2004 (first entry)

XX Human NOV18g.

(CURA-) CURAGEN CORP.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

XX Homo sapiens.

OS US2004038223-A1.

XX 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339268P.

PR 24-OCT-2001; 2001US-0343629P.

PR 01-NOV-2001; 2001US-0349575P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0373884P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.

PA (JULJ/) JU J.

PA (LILL/) LI L.

PA (GUOX/) GUO X.

PA (PATT/) PATTURAJAN M.

PA (SPYT/) SPYTEK K A.

PA (EDIN/) EDINGER S R.

PA (ELLE/) ELLERMAN K.

PA (MALY/) MALYANKAR U M.

PA (ORTT/) ORT T.

PA (GORM/) GORMAN L.

PA (ZERR/) ZERRHUSEN B D.

PA (ANDE/) ANDERSON D W.

PA (ZHON/) ZHONG M.

PA (CATT/) CATTERTON E.

PA (JIWW/) JI W.

PA (MILL/) MILLER C E.

PA (RST/) RASTELLI L.

PA (STON/) STONE D J.

PA (PENA/) PENNA C E A.

PA (SHEN/) SHENOY S G.

PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.

PA (LEAC/) LEACH M D.

PA (AGEE/) AGEE M L.

PA (BERG/) BERGHS C.  
 PA (DIP/) DIPIPPO V A.  
 PA (EISE/) EISEN A. E.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.  
 XX

PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerrhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoj SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Bergs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX

DR WPI; 2004-213931/20.

DR N-PSDB; ADN62907.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 104; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 224 AA;

Query Match 100.0%; Score 50; DB 8; Length 224;  
 Best Local Similarity 100.0%; Pred No. 2.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
 Db 193 GPLVCRGTL 201  
 |||||

RESULT 9

AAB98502

ID AAB98502 standard; protein; 225 AA.

XX

AC AAB98502;

DT 03-AUG-2001 (first entry)

DE Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.

XX Human; TAGD-15; cytostatic; vaccine; ovarian tumour; cancer;

KW

KW tumour antigen-derived gene 15; serine protease;  
KW Stratum Corneum Chymotryptic Enzyme; SCCE.

XX Homo sapiens.

FN WO200129056-A1.

XX 26-APR-2001.

PF 20-OCT-2000; 2000WO-US029095.

XX 20-OCT-1999; 99US-00421213.

XX (UYAR-) UNIV ARKANSAS.

PI O'brien TJ, Tanimoto H;

XX WPI; 2001-381031/40.

XX Novel extracellular serine protease, termed tumor antigen-derived gene 15  
PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,  
PT treatment, prevention of cancer, particularly breast, ovarian cancer.

XX Example 10; Fig 1; 130pp; English.

XX The present invention relates to human tumour antigen-derived gene 15  
CC (TADG-15) protein and coding sequence (see AAB23601 and AAB98500). TADG-  
CC 15 is an extracellular serine protease. It was found that TADG-15 is over  
CC residues that lack TADG-15 protease activity are useful for vaccinating  
CC an individual against TADG-15, having, suspected of having or at risk of  
CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic  
CC or therapeutic target in cancer. The present sequence was used in a  
XX sequence homology alignment with the catalytic domain of TADG-15

XX Sequence 225 AA;

Query Match 100.0%; Score 50; DB 4; Length 225;

Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 179 GPLVCRGTL 187

RESULT 10

ADA05742

XX ADA05742 standard; protein; 247 AA.

AC ADA05742;

XX 06-NOV-2003 (first entry)

XX Human NOV18f protein SEQ ID NO:102.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339286P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.

(CURA-) CURAGEN CORP.

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
Orr T, Gorman L, Zertusen BD, Anderson DW, Zhong M, Catterton E;  
Ji W, Miller CE, Rastelli L, Stone DU, Pena CBA, Shenoy SG;  
Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dippio VA;  
Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

WPI; 2003-381626/36.

N-PSDB; ADA05741.

New NOVX polypeptides and nucleic acids, useful for diagnosing,  
preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
pharmacogenomics.

Claim 1; Page 172; 586pp; English.

The present invention describes NOVX proteins, where X can be 1 to 55  
(e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
described above and a carrier; (2) a kit comprising, in one or more  
containers, the composition described above; (3) an isolated nucleic acid  
molecule which encodes a NOVX protein of the invention; (4) a vector  
comprising the nucleic acid molecule described above; (5) a cell  
comprising the above vector; (6) an antibody that immunospecifically  
binds to the polypeptide described above; (7) methods for determining the  
presence or amount of the above polypeptide or nucleic acid molecule in a  
sample; (8) methods for determining the presence of or predisposition to  
a disease associated with altered levels of expression of the above  
polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
method of identifying an agent that binds to the polypeptide described  
above; (10) a method for identifying a potential therapeutic agent for  
use in treating a pathology that is related to an aberrant expression or  
aberrant physiological interactions of the polypeptide; (11) a method of  
screening for a modulator of activity or of latency or predisposition to  
a pathology associated with the polypeptide; (12) a method for modulating  
the activity of the polypeptide described above; (13) methods of treating  
or preventing a pathology associated with the above polypeptide in a  
mammal; and (14) a method for producing the above polypeptide. NOVX  
sequences have antidiabetic, anorectic, antibacterial, virucide,  
immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
and antilipemic activities, and can be used in gene therapy. The  
polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 247 AA;  
 Query Match 100.0%; Score 50; DB 6; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
 Db 193 GPLVCRGTL 201

RESULT 11  
 ADN62906  
 ID ADN62906 standard; protein; 247 AA.

XX AC ADN62906;

XX DT 01-JUL-2004 (first entry)  
 XX DE Human NOV18f.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

OS Homo sapiens.

XX PN US2004038223-A1.

XX PD 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 12-OCT-2001; 2001US-0328056P.

PR 15-OCT-2001; 2001US-0328849P.

PR 17-OCT-2001; 2001US-0329414P.

PR 18-OCT-2001; 2001US-0330142P.

PR 22-OCT-2001; 2001US-0330309P.

PR 24-OCT-2001; 2001US-0341058P.

PR 29-OCT-2001; 2001US-0343629P.

PR 01-NOV-2001; 2001US-0349575P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0373884P.

PR 25-JUN-2002; 2002US-0391335P.

PA (SMIT//) SMITHSON G.

PA (MILL//) MILLET I.

PA (PEYM//) PEYMAN J A.

PA (KEKU//) KEKUDA R.

PA (JUUJ//) JU J.

PA (LILL//) LI L.

PA (GUOX//) GUO X.

PA (PATT//) PATTURAJAN M.

PA (SPYT//) SPYTEK K A.

PA (EDIN//) EDINGER S R.

PA (ELLE//) ELLERMAN K.

PA (MALY//) MALYANKAR U M.

PA (ORTT//) ORT T.

PA (GORM//) GORMAN L.

PA (ZERH//) ZERHUSEN B D.

PA (ANDE//) ANDERSON D W.

PA (ZHON//) ZHONG M.

PA (CATT//) CATTERTON E.

PA (JIWW//) JI W.

PA (MILL//) MILLER C E.

PA (RST//) RASTELLI L.

PA (STON//) STONE D J.

PA (PENA//) PENA C E A.

PA (SHEN//) SHENOY S G.

PA (SHIM//) SHIMKETS R A.

PA (ROTH//) ROTHENBERG M E.

PA (LEAC//) LEACH M D.

PA (AGEE//) AGEE M L.

PA (BERG//) BERGHS C.

PA (DIP//) DIPIPPO V A.

PA (EISE//) EISEN A.

PA (GANG//) GANGOLLI E A.

PA (RIEG//) RIEGER D K.

PA (SPAD//) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2004-213931/20.

DR N-PSDB; ADN62905.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,

PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 102; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.

CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or

CC treat a medical condition in human related to the aberrant expression and

CC activity of NOVX polypeptides. For example, NOVX polypeptides and

CC polynucleotides may be used to treat disorders associated with decreased

CC expression or activity of NOVX by supplementing the patient our

CC production or to rectify mutations. Conversely, antisense NA molecules

CC may be administered to down regulate expression of NOVX polypeptides by

CC binding with the cells own genes and preventing their expression. NOVX

CC polynucleotides and complementary sequences may also be used as DNA

CC probes in diagnostic assays to detect and quantify the presence of

CC similar sequences in samples, and so which patients may be in need of

CC restorative therapy. NOVX polypeptides may also be used as antigens in

CC the production of antibodies and in assays to identify modulators

CC (agonists and antagonists) of the expression and activity of NOVX. The

CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be

CC used to modulate NOVX polynucleotide expression and activity of NOVX

CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as

CC diagnostic agents for detecting the presence of NOVX in samples. NOVX

CC polypeptides and polynucleotides may be used in this way to prevent,

CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious

CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 247 AA;  
  
Query Match 100.0%; Score 50; DB 8; Length 247;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GPLVCRGTL 9  
Db |||||  
193 GPLVCRGTL 201  
  
RESULT 12  
ADA05732  
ID ADA05732 standard; protein; 250 AA.  
XX  
AC ADA05732;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV18a protein SEQ ID NO:92.  
XX  
KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipaemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
PF  
PP  
PX 02-OCT-2002; 2002WO-US031373.  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 09-OCT-2001; 2001US-0327443P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 12-OCT-2001; 2001US-0328056P.  
PR 15-OCT-2001; 2001US-0328849P.  
PR 17-OCT-2001; 2001US-0329414P.  
PR 18-OCT-2001; 2001US-0330142P.  
PR 22-OCT-2001; 2001US-0330309P.  
PR 24-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 29-OCT-2001; 2001US-0343629P.  
PR 01-NOV-2001; 2001US-0349575P.  
PR 17-APR-2002; 2002US-0346357P.  
PR 19-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.  
XX (CURA-) CURAGEN CORP.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiIppio VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2003-381626/36.  
DR N-PSDB; ADA05731.  
XX  
PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Claim 1; Page 169-170; 586pp; English.  
XX  
CC The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
CC and antilipaemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and the  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
XX present invention.  
XX Sequence 250 AA;  
SQ  
  
Query Match 100.0%; Score 50; DB 6; Length 250;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GPLVCRGTL 9  
Db |||||  
204 GPLVCRGTL 212  
  
RESULT 13  
ADN62896  
ID ADN62896 standard; protein; 250 AA.  
XX  
AC ADN62896;  
XX



DT XX 01-JUL-2004 (first entry)  
DE XX Human NOV18a.  
KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.  
XX  
OS Homo sapiens.  
XX  
XX US2004038223-A1.  
XX  
XX 26-FEB-2004.  
XX  
XX 01-OCT-2002; 2002US-00262511.  
XX  
XX 02-OCT-2001; 2001US-0326483P.  
XX 05-OCT-2001; 2001US-0327435P.  
XX 09-OCT-2001; 2001US-0327449P.  
XX 09-OCT-2001; 2001US-0327917P.  
XX 09-OCT-2001; 2001US-0328029P.  
XX 09-OCT-2001; 2001US-0328044P.  
XX 09-OCT-2001; 2001US-0328056P.  
XX 12-OCT-2001; 2001US-0328849P.  
XX 15-OCT-2001; 2001US-0329414P.  
XX 17-OCT-2001; 2001US-0330142P.  
XX 18-OCT-2001; 2001US-0330309P.  
XX 22-OCT-2001; 2001US-0341058P.  
XX 24-OCT-2001; 2001US-0332668P.  
XX 24-OCT-2001; 2001US-0343629P.  
XX 29-OCT-2001; 2001US-0349575P.  
XX 01-NOV-2001; 2001US-0346357P.  
XX 17-APR-2002; 2002US-0373260P.  
XX 19-APR-2002; 2002US-0373815P.  
XX 19-APR-2002; 2002US-0373817P.  
XX 19-APR-2002; 2002US-0373828P.  
XX 19-APR-2002; 2002US-0373884P.  
XX 22-APR-2002; 2002US-0374977P.  
XX 16-MAY-2002; 2002US-0381037P.  
XX 16-MAY-2002; 2002US-0381038P.  
XX 17-MAY-2002; 2002US-0381642P.  
XX 28-MAY-2002; 2002US-0383656P.  
XX 29-MAY-2002; 2002US-0383831P.  
XX 25-JUN-2002; 2002US-0391335P.  
XX (SMT/) SMITHSON G.  
PA (MILL/) MILLET I.  
PA (PEYM/) PEYMAN J A.  
PA (KEKU/) KEKUDA R.  
PA (JUJJ/) JU J.  
PA (LILL/) LI L.  
PA (GUOX/) GUO X.  
PA (PATT/) PATTURAJAN M.  
PA (SPYT/) SPYTEK K A.  
PA (EDIN/) EDINGER S R.  
PA (ELLE/) ELLERMAN K.  
PA (MALY/) MALYANKAR U M.  
PA (ORTT/) ORT T.  
PA (GORM/) GORMAN L.  
PA (ZERR/) ZERHUSEN B D.  
PA (ANDE/) ANDERSON D W.  
PA (ZHON/) ZHONG M.  
PA (CATT/) CATTERTON E.  
PA (JLWW/) JI W.  
PA (MILL/) MILLER C E.  
PA (RAST/) RASTELLI L.  
PA (STON/) STONE D J.  
PA (PENA/) PENNA C E A.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.  
PA (LEAC/) LEACH M D.  
PA (AGEE/) AGE E M L.  
PA (BERG/) BERGHS C.  
PA (DIP1/) DIPIPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
XX WPI: 2004-213931/20.  
DR N-PSDB; ADN62895.  
XX  
XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
XX Claim 1; SEQ ID NO 92; 395pp; English.  
XX  
XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
XX represents the amino acid sequence of a human NOVX protein.

XX Sequence 250 AA;

Query Match 100.0%; Score 50; DB 8; Length 250;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 204 GPLVCRGTL 212

RESULT 14

ADA05734  
ID ADA05734 standard; protein; 252 AA.

XX  
AC ADA05734;

XX 06-NOV-2003 (first entry)

XX

DE Human NOV18b protein SEQ ID NO:94.  
XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cyostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO2003029424-A2.  
XX  
XX 10-APR-2003.  
XX  
XX 02-OCT-2002; 2002WO-US031373.  
XX  
XX 02-OCT-2001; 2001US-0326483P.  
XX 05-OCT-2001; 2001US-0327435P.  
XX 05-OCT-2001; 2001US-0327449P.  
XX 09-OCT-2001; 2001US-0327917P.  
XX 09-OCT-2001; 2001US-0328029P.  
XX 09-OCT-2001; 2001US-0328044P.  
XX 09-OCT-2001; 2001US-0328056P.  
XX 09-OCT-2001; 2001US-0328059P.  
XX 12-OCT-2001; 2001US-0328849P.  
XX 15-OCT-2001; 2001US-0329141P.  
XX 17-OCT-2001; 2001US-0330142P.  
XX 18-OCT-2001; 2001US-0330309P.  
XX 22-OCT-2001; 2001US-0341058P.  
XX 24-OCT-2001; 2001US-0339266P.  
XX 24-OCT-2001; 2001US-0343629P.  
XX 29-OCT-2001; 2001US-0349575P.  
XX 01-NOV-2001; 2001US-0346357P.  
XX 17-APR-2002; 2002US-0373260P.  
XX 19-APR-2002; 2002US-0373815P.  
XX 19-APR-2002; 2002US-0373817P.  
XX 19-APR-2002; 2002US-0373826P.  
XX 19-APR-2002; 2002US-0373884P.  
XX 22-APR-2002; 2002US-0374977P.  
XX 16-MAY-2002; 2002US-0381037P.  
XX 16-MAY-2002; 2002US-0381038P.  
XX 16-MAY-2002; 2002US-0381042P.  
XX 17-MAY-2002; 2002US-0381648P.  
XX 28-MAY-2002; 2002US-0383656P.  
XX 29-MAY-2002; 2002US-0383831P.  
XX 25-JUN-2002; 2002US-0391335P.  
XX 01-OCT-2002; 2002US-00262511.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerkhus BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shinkets RA, Rotherberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
XX WPI; 2003-381626/36.  
DR N-P5DB; ADA05733.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX Claim 1; Page 170; 586pp; English.  
XX  
XX The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell

CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cyostatic, nootropic, neuroprotective, antiparkinsonian  
CC and antilipemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer, a  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX  
XX SQ Sequence 252 AA;

Query Match 100.0%; Score 50; DB 6; Length 252;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 206 GPLVCRGTL 214  
|||||

RESULT 15

ADN62898  
ID ADN62898 standard; protein; 252 AA.

AC ADN62898;

DT 01-JUL-2004 (first entry)

DE Human NOV18b.

KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
KW wasting disorder.

OS Homo sapiens.

XX US2004038223-A1.

XX 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 05-OCT-2001; 2001US-0327449P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328849P.

XX 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX  
 PA (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPVT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIP/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.  
 XX  
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 DR WPI; 2004-213931/20.  
 DR N-PSDB; ADN62897.  
 XX  
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 XX  
 PS Claim 1; SEQ ID NO 94; 395pp; English.  
 XX  
 CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or

CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX

SQ Sequence 252 AA;

Query Match 100.0%; Score 50; DB 8; Length 252;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 206 GPLVCRGTL 214

Search completed: July 13, 2005, 17:19:55

Job time : 77.6667 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083A-80

Perfect score: 50

Sequence: 1 GPLVCRGTL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.\*

1: Pirl.\*

2: Pirl2.\*

3: Pirl3.\*

4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	253	1 A53968	serine proteinase
2	40	80.0	261	2 A32297	semenogelase (EC 3
3	40	80.0	261	2 A29586	tissue kallikrein
4	40	80.0	262	2 T35999	probable aminoglyc
5	39	78.0	155	2 G72548	hypothetical prote
6	39	78.0	265	1 KORTP	tissue kallikrein
7	38	76.0	246	1 DBHU	complement factor
8	38	76.0	261	1 S35711	semenogelase (EC 3
9	37	74.0	46	2 I49416	glandular kallikre
10	37	74.0	96	2 A05308	tissue kallikrein
11	37	74.0	104	2 S15395	tissue kallikrein
12	37	74.0	156	2 B32863	tissue kallikrein
13	37	74.0	188	2 B32340	tissue kallikrein
14	37	74.0	225	2 S45356	probable serine pr
15	37	74.0	239	2 A27207	tissue kallikrein
16	37	74.0	240	2 S39047	trypsin (EC 3.4.21
17	37	74.0	241	2 S39048	trypsin (EC 3.4.21
18	37	74.0	244	2 A44284	tissue kallikrein
19	37	74.0	250	2 S15685	kallikrein, glandu
20	37	74.0	254	2 S35585	chymotrypsin-like
21	37	74.0	258	2 I56220	trypsin 2 - rat
22	37	74.0	259	1 KORTTN	tonin (EC 3.4.21.-
23	37	74.0	259	2 A29746	tissue kallikrein
24	37	74.0	259	2 D23863	tissue kallikrein
25	37	74.0	260	2 I56559	neurosin - mouse
26	37	74.0	261	1 KQMS1	tissue kallikrein
27	37	74.0	261	1 NGMSG	7S nerve growth fa
28	37	74.0	261	1 TRMSG5	tissue kallikrein
29	37	74.0	261	2 S45303	tissue kallikrein

30	37	74.0	261	2 A28062	gamma-renin (EC 3.
31	37	74.0	261	2 A25606	tissue kallikrein
32	37	74.0	261	2 S01971	tissue kallikrein
33	37	74.0	261	2 A29745	tissue kallikrein
34	37	74.0	261	2 A34079	tissue kallikrein
35	37	74.0	263	2 S15686	tissue kallikrein
36	37	74.0	264	2 S65663	granzyme 3 (EC 3.4
37	37	74.0	599	2 T12994	riboflavin biosynt
38	36	72.0	66	2 I52972	kallikrein - mouse
39	36	72.0	73	2 S44462	elastase (EC 3.4.2
40	36	72.0	149	1 KQMSM	tissue kallikrein
41	36	72.0	226	2 S69370	duodenase - bovine
42	36	72.0	229	1 TRDES	trypsin (EC 3.4.21
43	36	72.0	231	1 TRPGR	trypsin (EC 3.4.21
44	36	72.0	231	2 S31778	trypsin (EC 3.4.21
45	36	72.0	236	1 A32121	snake venom factor

ALIGNMENTS

RESULT 1

A53968  
serine proteinase SCCE precursor - human  
N:Alternate names: stratum corneum chymotryptic enzyme  
C:Species: Homo sapiens (man)  
C:Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: A53968  
R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.  
J. Biol. Chem. 269, 19420-19426, 1994  
A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme  
A:Reference number: A53968; MUID:94308225; PMID:8034709  
A:Accession: A53968  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-253 <HAN>  
A:Cross-references: UNIPROT:P49862; GB:L33404; NID:9521214; PIDN:AAC37551.1; PID:9532504  
C:Genetics:  
A:Gene: GDB:PRSS6; SCCE  
A:Cross-references: GDB:377730  
A:Map position: 7q35-7q35  
C:Superfamily: trypsin; trypsin homology  
F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 50; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
| | | | | | | | |  
Db 207 GPLVCRGTL 215

RESULT 2

A32297  
semenogelase (EC 3.4.21.77) precursor [validated] - human  
N:Alternate names: gamma-senoinoprotein; P-30 antigen; prostate-specific antigen; prostate  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C:Accession: A32297; A32423; S03604; S05468; S05467; A32546; S02239; A26757; C31567; A315  
R:Riegman, P.H.J.; Vlieestra, R.J.; van der Korput, J.A.G.M.; Romijn, J.C.; Trapman, J.  
Biochem. Biophys. Res. Commun. 159, 95-102, 1989  
A:Title: Characterization of the prostate-specific antigen gene: a novel human kallikrein  
A:Reference number: A32297; MUID:89165891; PMID:2466464  
A:Accession: A32297  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-261 <RIE>  
A:Cross-references: UNIPROT:P07288; GB:M24543  
R:Lundwall, A.  
Biochem. Biophys. Res. Commun. 161, 1151-1159, 1989  
A:Title: Characterization of the gene for prostate-specific antigen, a human glandular ke  
A:Reference number: A32423; MUID:89302090; PMID:2472789

A:Accession: A32423  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-261 <LUN>  
A:Cross-references: GB:M27274; NID:g190552; PIDN:AAA60192.1; PID:g190553  
R:Digby, M.; Zhang, X.Y.; Richards, R.I.  
Nucleic Acids Res. 17, 2137, 1989  
A:Title: Human prostate specific antigen (PSA) gene: structure and linkage to the kallikrein gene  
A:Reference number: S03604; MUID:89183632; PMID:2467258  
A:Accession: S03604  
A:Molecule type: DNA  
A:Residues: 1-261 <DIG>  
A:Cross-references: EMBL:X13940  
R:Klobeck, H.G.; Combratio, G.; Schulz, P.; Arbusow, V.; Fittler, F.  
submitted to the EMBL Data Library, May 1989  
A:Reference number: S05468  
A:Accession: S05468  
A:Molecule type: DNA  
A:Residues: 1-261 <KL1>  
A:Cross-references: EMBL:X14810; NID:g35732; PIDN:CAA32915.1; PID:g296671  
R:Klobeck, H.G.; Combratio, G.; Schulz, P.; Arbusow, V.; Fittler, F.  
Nucleic Acids Res. 17, 3981, 1989  
A:Title: Genomic sequence of human prostate specific antigen (PSA).  
A:Reference number: S05467; MUID:89282407; PMID:2471958  
A:Accession: S05467  
A:Molecule type: DNA  
A:Residues: 1-29 <KL2>  
A:Cross-references: EMBL:X14810  
R:Henttu, P.; Viikho, P.  
Biochem. Biophys. Res. Commun. 160, 903-910, 1989  
A:Title: cDNA coding for the entire human prostate specific antigen shows high homologie  
A:Reference number: A32546; MUID:89246551; PMID:2470373  
A:Accession: A32546  
A:Molecule type: mRNA  
A:Residues: 1-72, 'T', 74-85, 'I', 87-174, 'P', 176-183, 'Q', 185-259, 'D', 261 <HEN>  
A:Cross-references: GB:M26663  
R:Schulz, P.; Stucka, R.; Feldmann, H.; Combratio, G.; Klobeck, H.G.; Fittler, F.  
Nucleic Acids Res. 16, 6226, 1988  
A:Title: Sequence of a cDNA clone encompassing the complete mature human prostate specific antigen  
A:Reference number: S02239; MUID:8828366; PMID:2456523  
A:Accession: S02239  
A:Molecule type: mRNA  
A:Residues: 17-63, 'T', 65-135, 'M', 137-261 <SCH>  
A:Cross-references: EMBL:X07730  
R:Lundwall, A.; Lilja, H.  
FEBS Lett. 214, 317-322, 1987  
A:Title: Molecular cloning of human prostate specific antigen cDNA.  
A:Reference number: A26757; MUID:87190978; PMID:2436946  
A:Accession: A26757  
A:Molecule type: mRNA  
A:Residues: 5-261 <LU2>  
R:Riegman, P.H.J.; Klaassen, P.; van der Korput, J.A.G.M.; Romijn, J.C.; Trapman, J.  
Biochem. Biophys. Res. Commun. 155, 181-188, 1988  
A:Title: Molecular cloning and characterization of novel prostate antigen cDNA's.  
A:Reference number: A90144; MUID:88326297; PMID:2458104  
A:Accession: C31567  
A:Molecule type: mRNA  
A:Residues: 5-261 <RT2>  
A:Cross-references: GB:M21895; NID:g189523; PIDN:AAA59995.1; PID:g189524  
A:Note: the authors translated the codon GGC for residue 28 as Arg and TGG for residue 29 as Arg  
A:Accession: A31567  
A:Molecule type: mRNA  
A:Residues: 5-210, 'WVILITELTMPA', 223, 'PMVLHGLV', 233, 'WRGGV' <RT3>  
A:Cross-references: GB:M21896; NID:g189525; PIDN:AAA59996.1; PID:g189526  
A:Note: the authors translated the codon GGC for residue 28 as Arg and TGG for residue 29 as Arg  
A:Accession: B31567  
A:Molecule type: mRNA  
A:Residues: 5-164, 'CTPGPDGAAGSPDAW' <RT4>  
A:Cross-references: GB:M21897; NID:g189529; PIDN:AAA59997.1; PID:g189530  
A:Note: the authors translated the codon GGC for residue 28 as Arg, TGG for residue 29 as Arg  
R:Schaller, J.; Akiyama, K.; Tauda, R.; Hara, M.; Marti, T.; Rickli, E.B.  
Eur. J. Biochem. 170, 111-120, 1987  
A:Title: Isolation, characterization and amino-acid sequence of gamma-seminoprotein, a

A:Reference number: S00232; MUID:88082806; PMID:3691515  
A:Accession: S00232  
A:Molecule type: protein  
A:Residues: 25-261 <SC2>  
R:Watt, K.W.K.; Lee, P.J.; M'Timkulu, T.; Chan, W.P.; Looor, R.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3166-3170, 1986  
A:Title: Human prostate-specific antigen: structural and functional similarity with serin  
A:Reference number: A23937; MUID:86205857; PMID:2422647  
A:Accession: A23937  
A:Molecule type: protein  
A:Residues: 25-93, 'T', 95-164, 'HL', 166, 'YDQM', 169-174, 'Q', 176-261 <WAT>  
R:Moreno, J.M.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: G07735  
A:Accession: G01551  
A:Status: translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-261 <MOR>  
A:Cross-references: EMBL:U17040; NID:g595945; PIDN:AAA56764.1; PID:g595946  
R:Monne, M.; Croce, C.M.; Yu, H.; Diamandis, E.P.  
Cancer Res. 54, 6344-6347, 1994  
A:Title: Molecular characterization of prostate-specific antigen messenger RNA expressed  
A:Reference number: I52712; MUID:95079406; PMID:7527295  
A:Accession: I52712  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 'PQAVLHPESCVT', 1-176 <RES>  
A:Cross-references: GB:S75755; NID:9861469; PIDN:AAD14185.1; PID:g4261885  
R:Christenson, A.; Lilja, H.  
Eur. J. Biochem. 220, 45-53, 1994  
A:Title: Complex formation between protein C inhibitor and prostate-specific antigen in v  
A:Reference number: S41212; MUID:94164172; PMID:7509746  
A:Accession: S41212  
A:Molecule type: protein  
A:Residues: 25-30, 'X', 32-49 <CHR>  
C:Comment: This enzyme preferentially cleaves after tyrosine residues.  
C:Genetics:  
A:Gene: GDB:APS; PSA  
A:Cross-references: GDB:119695; OMIM:176820  
A:Map position: 19q13.3-19q13.3  
A:Introns: 16/1; 69/2; 165/1; 210/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: glycoprotein; hydrolase; prostate; serine proteinase  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-24/Domain: propeptide #status predicted <PRO>  
F:25-261/Product: semogelase #status experimental <WAT>  
F:25-253/Domain: trypsin homology <TRY>  
F:65,120,213/Active site: His, Asp, Ser #status predicted  
Query Match 80.0%; Score 40; DB 1; Length 261;  
Best Local Similarity 77.8%; Pred. No. 6.4;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 GPLVCRGTL 9  
Db 215 GPLVCGVL 223  
RESULT 3  
A29586  
tissue kallikrein (EC 3.4.21.35) hGK-1 precursor - human  
A:Alternate names: glandular kallikrein  
C:Species: Homo sapiens (man)  
C:Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 09-Jul-2004  
C:Accession: A29586  
R:Schedlich, L.J.; Bennetts, B.H.; Morris, B.J.  
DNA 6, 429-437, 1987  
A:Title: Primary structure of a human glandular kallikrein gene.  
A:Reference number: A29586; MUID:88054467; PMID:2824146  
A:Accession: A29586  
A:Molecule type: DNA  
A:Residues: 1-261 <SCH>  
A:Cross-references: UNIPROT:P20151; GB:M18157; NID:g186640; PIDN:AAA74454.1; PID:g386842

A>Note: the authors translated the codon TAC for residue 43 as Trp

C:Genetics: 16/1; 69/2; 165/1; 210/3  
 A:Introns: 16/1; 69/2; 165/1; 210/3  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:25-253/Domain: trypsin homology <TRY>  
 F:65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 80.0%; Score 40; DB 2; Length 261;  
 Best Local Similarity 77.8%; Pred. No. 6.4;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0;  
 Gaps 0;

QY 1 GPLVCRGTL 9  
 |||||  
 DB 215 GPLVNGVL 223

RESULT 4  
 T35999  
 Probable aminoglycoside acetyltransferase - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T35999  
 R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, July 1999  
 A:Reference number: Z1574  
 A:Accession: T35999  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-262 <SEE>  
 A:Cross-references: UNIPROT:Q9XAC9; EMBL:AL096839; PIDN:CA850752.1; GSPDB:GN00070; SCOPED  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 C:Superfamily: Escherichia coli aminoglycoside N3'-acetyltransferase

Query Match 80.0%; Score 40; DB 2; Length 262;  
 Best Local Similarity 88.9%; Pred. No. 6.5;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9  
 |||||  
 DB 8 GPLVTRGTL 16

RESULT 5  
 G72548  
 Hypothetical protein APE1675 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: G72548  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: G72548  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-155 <KAW>  
 A:Cross-references: UNIPROT:Q9YBC2; DDBJ:AP000062; NID:g5105244; PIDN:BAA80676.1; PID:d1  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE1675

Query Match 78.0%; Score 39; DB 2; Length 155;  
 Best Local Similarity 77.8%; Pred. No. 6.2;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9  
 |||||  
 DB 23 GPLITRGTL 31

# RESULT 6

## KORTP

tissue kallikrein (EC 3.4.21.35) precursor - rat  
 N:Alternate names: glandular kallikrein; Kininogenin; true tissue kallikrein  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 05-Apr-1983 #sequence\_revision 05-Apr-1983 #text\_change 09-Jul-2004  
 C:Accession: A00944; A41429; A25137; JX0073; A23863; A33359  
 R:Swift, G.H.; Dagorn, J.C.; Ashley, P.L.; Cummings, S.W.; MacDonald, R.J.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 7263-7267, 1982  
 A:Title: Rat pancreatic kallikrein mRNA: nucleotide sequence and amino acid sequence of t  
 A:Reference number: A00944; MUID:83117659; PMID:6961406  
 A:Accession: A00944  
 A:Molecule type: mRNA  
 A:Residues: 1-265 <SWI>  
 A:Cross-references: UNIPROT:P00758  
 R:Kato, H.; Nakanishi, E.; Enjoji, K.; Hayashi, I.; Oh-Ishi, S.; Iwanaga, S.  
 J. Biochem. 102, 1389-1404, 1987  
 A:Title: Characterization of serine proteinases isolated from rat submaxillary gland: wit  
 A:Reference number: A41429; MUID:88198057; PMID:3482210  
 A:Accession: A41429  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 29-53 'X', 55-87 <KAT>  
 R:Gerald, W.L.; Chao, J.; Chao, L.  
 Biochim. Biophys. Acta 866, 1-14, 1986  
 A:Title: Immunological identification of rat tissue kallikrein cDNA and characterization  
 A:Reference number: A25137; MUID:86131678; PMID:3004582  
 A:Accession: A25137  
 A:Molecule type: mRNA  
 A:Residues: 115-265 <GER>  
 R:Inoue, H.; Fukui, K.; Miyake, Y.  
 J. Biochem. 105, 834-840, 1989  
 A:Title: Identification and structure of the rat true tissue kallikrein gene expressed in  
 A:Reference number: JX0073; MUID:89327211; PMID:2753879  
 A:Accession: JX0073  
 A:Molecule type: DNA  
 A:Residues: 1-265 <INO>  
 A:Cross-references: GB:D00448; NID:g220792; PIDN:BAA00346.1; PID:g220794  
 A:Experimental source: kidney  
 R:Ashley, P.L.; MacDonald, R.J.  
 Biochemistry 24, 4512-4520, 1985  
 A:Title: Kallikrein-related mRNAs of the rat submaxillary gland: nucleotide sequences of  
 A:Reference number: A23863; MUID:86051477; PMID:2998455  
 A:Accession: A23863  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-265 <ASH>  
 A:Cross-references: GB:M11563; NID:g205029; PIDN:AAA41464.1; PID:g205030  
 A:Experimental source: submaxillary gland  
 R:Wines, D.R.; Brady, J.M.; Pritchett, D.B.; Roberts, J.L.; MacDonald, R.J.  
 J. Biol. Chem. 264, 7653-7662, 1989  
 A:Title: Organization and expression of the rat kallikrein gene family.  
 A:Reference number: A33359; MUID:89214217; PMID:2708383  
 A:Accession: A33359  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 5-265 <WIN>  
 A:Cross-references: GB:M23874; GB:J04701; GB:M23875; GB:M23876; NID:g205007; PIDN:AAA4146  
 C:Comment: The kallikreins liberate lysyl-bradykinin, a vasoactive decapeptide, from kin  
 C:Comment: The protein presumably assumes the two-chain form by cleavage between residu  
 C:Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release l  
 C:Genetics:  
 A:Introns: 20/1; 73/2; 169/1; 214/3  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; pancreas; serine proteinase; zymogen  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-28/Domain: activation peptide #status predicted <APT>  
 F:29-265/Product: tissue kallikrein, pancreatic #status predicted <MPT>  
 F:29-257/Domain: trypsin homology <TRY>  
 F:35-177, 54-70, 156-223, 188-202, 213-238/Disulfide bonds: #status predicted  
 F:169,124,217/Active site: His, Asp, Ser #status predicted

Query Match 78.0%; Score 39; DB 1; Length 265;  
Best Local Similarity 66.7%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||:|:|  
Db 219 GPLICNGVL 227

RESULT 7  
DBHU  
complement factor D (EC 3.4.21.46) precursor [validated] - human (fragment)  
N;Alternate names: adipsin; C3 convertase activator  
C;Species: Homo sapiens (man)  
C;Date: 28-Aug-1995 #sequence revision 31-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: A40197; A00936; A60571; S66645  
R;White, R.T.; Damm, D.; Hancock, N.; Rosen, B.S.; Lowell, B.B.; Usher, P.; Flier, J.S.; J. Biol. Chem. 267, 9210-9213, 1992  
A;Title: Human adipsin is identical to complement factor D and is expressed at high level  
A;Reference number: A40197; MUID:92250520; PMID:1374388  
A;Accession: A40197  
A;Molecule type: mRNA  
A;Residues: 1-246 <WHI>  
A;Cross-references: UNIPROT:P00746; GB:M84526  
R;Niemann, M.A.; Bhow, A.S.; Bennett, J.C.; Volanakis, J.E. Biochemistry 23, 2482-2486, 1984  
A;Title: Amino acid sequence of human D of the alternative complement pathway.  
A;Reference number: A00936; MUID:85000441; PMID:6383466  
A;Accession: A00936  
A;Molecule type: protein  
A;Residues: 19-44,'G',46-51,'Q',53-75,'TH',78,'P',80-83,'XXXITIE',90-172,86-91,185-235,'  
A;Note: a few residues were assigned from the previously published sequence of Reid et al  
R;Miyata, T.; Oda, O.; Inagi, R.; Sugiyama, S.; Miyama, A.; Maeda, K.; Nakashima, I.; Y Mol. Immunol. 27, 637-644, 1990  
A;Title: Molecular and functional identification and purification of complement component  
A;Reference number: A60571; MUID:90370044; PMID:2395435  
A;Accession: A60571  
A;Molecule type: protein  
A;Residues: 19-20,'XX',23-27,'XX',30-31,'XX',34,'X',36-40 <MIY>  
R;Baake, N.; Holtkamp, U.; Hoerl, W.H.; Tschesche, H. FEBS Lett. 371, 300-302, 1995  
A;Title: Inhibition of degranulation of human polymorphonuclear leukocytes by complement  
A;Reference number: S66645; MUID:96013156; PMID:7556615  
A;Accession: S66645  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 19-44,'C',46-48 <BAL>  
C;Comment: Factor D cleaves factor B when the latter is complexed with factor C3b, activ  
C;Genetics:  
A;Gene: GDB:DF  
A;Cross-references: GDB:132645; OMIM:134350  
A;Map position: Xpter-Xqter  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-246/Product: complement factor D (fragment) #status experimental <MAT>  
F;19-241/Domain: trypsin homology <TRY>  
F;44-60,141-207,172-188,197-222/Disulfide bonds: #status predicted  
F;59,105,201/Active site: His, Asp, Ser #status predicted

Query Match 76.0%; Score 38; DB 1; Length 246;  
Best Local Similarity 77.8%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||:|:|  
Db 203 GPLVCGVL 211

RESULT 8  
S35711  
semenogelase (EC 3.4.21.77) precursor - rhesus macaque

N;Alternate names: gamma-seminoprotein; P-30 antigen; prostate-specific antigen; prostate  
C;Species: Macaca mulatta (rhesus macaque)  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C;Accession: S35711; S34239  
R;Gauthier, E.R.; Chapdelaine, P.; Tremblay, R.R.; Dube, J.Y. Biochim. Biophys. Acta 1174, 207-210, 1993  
A;Title: Characterization of rhesus monkey prostate specific antigen cDNA.  
A;Reference number: S35711; MUID:93363642; PMID:7689340  
A;Accession: S35711  
A;Molecule type: mRNA  
A;Residues: 1-261 <GAU>  
A;Cross-references: UNIPROT:P33619; EMBL:X73560; NID:G311843; PIDN:CAA51957.1; PID:G31184  
C;Comment: This enzyme preferentially cleaves after tyrosine residues.  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: glycoprotein; hydrolase; prostate; serine proteinase  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-24/Domain: propeptide #status predicted <PRO>  
F;25-261/Product: semenogelase #status predicted <MAT>  
F;25-253/Domain: trypsin homology <TRY>  
F;65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 76.0%; Score 38; DB 1; Length 261;  
Best Local Similarity 77.8%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||:|:|  
Db 215 GPLVCDGVL 223

RESULT 9  
I49416  
glandular kallikrein - western wild mouse (fragment)  
C;Species: Mus spretus (western wild mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49416  
R;Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H. Mamm. Genome 5, 349-355, 1994  
A;Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
A;Reference number: I48934; MUID:94319082; PMID:8043949  
A;Accession: I49416  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-46 <RES>  
A;Cross-references: UNIPROT:Q62540; EMBL:U05716; NID:G497047; PIDN:AAB60472.1; PID:G49704  
C;Superfamily: trypsin; trypsin homology

Query Match 74.0%; Score 37; DB 2; Length 46;  
Best Local Similarity 66.7%; Pred. No. 4.8;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||:|:|  
Db 1 GPLICDGVL 9

RESULT 10  
A05308  
tissue kallikrein (EC 3.4.21.35), submandibular MGK-2 - mouse (fragment)  
N;Alternate names: Glandular kallikrein  
C;Species: Mus musculus (house mouse)  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 22-Jun-1999  
C;Accession: A05308  
R;Mason, A.J.; Evans, B.A.; Cox, D.R.; Shine, J.; Richards, R.I. Nature 303, 300-307, 1983  
A;Title: Structure of mouse kallikrein gene family suggests a role in specific processing  
A;Reference number: A00941; MUID:83219214; PMID:6602295  
A;Accession: A05308  
A;Molecule type: DNA  
A;Residues: 1-96 <MAS>  
A;Cross-references: GB:V00829; NID:G52775; PIDN:CAA24212.1; PID:G388523  
A;Experimental source: Quakenbush inbred strain  
A;Note: this sequence has been translated from two exons (11-147 and 522-674) located up



C;Comment: Glandular kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release  
C;Genetics:  
A;Gene: mGK-2  
A;Map position: 7  
A;Introns: 45/3  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: glycoprotein; hydrolase; saliva; serine proteinase; submandibular gland  
F;1-88/Domain: trypsin homology (fragment) <TRY>  
  
Query Match 74.0%; Score 37; DB 2; Length 96;  
Best Local Similarity 66.7%; Pred. No. 9.3;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 GPLVCRGTL 9  
Db 50 GPLICDGV 58  
		:
		:
		:
		:
  
RESULT 11  
S15395  
tissue kallikrein-related proteinase (EC 3.4.21.-) k10 - rat (fragments)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 18-Feb-1994 #sequence\_revision 19-Jul-1996 #text\_change 17-Oct-1997  
C;Accession: S15395; S15448  
R;Gutman, N.; Elmoujahed, A.; Brillard, M.; Monegier du Sorbier, B.; Gauthier, F.  
Eur. J. Biochem. 197, 425-429, 1991  
A;Title: Microheterogeneity of rat submaxillary gland kallikrein k10, a member of the ka  
A;Reference number: S15395; MUID:91224135; PMID:2026164  
A;Accession: S15395  
A;Molecule type: protein  
A;Residues: 1-23;24-51:52-104 <GUT>  
A;Experimental source: submaxillary gland  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-104/Domain: trypsin homology (fragments) <TRY>  
F;1-23/Product: tissue kallikrein-related protein k10 light chain (fragment) #status ex  
F;24-51/Product: tissue kallikrein-related protein k10 heavy chain (fragment) #status ex  
F;52-104/Product: tissue kallikrein-related protein k10 heavy chain (fragment) #status e  
  
Query Match 74.0%; Score 37; DB 2; Length 104;  
Best Local Similarity 66.7%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 GPLVCRGTL 9  
Db 70 GPLICDGV 78  
		:
		:
		:
		:
  
RESULT 12  
B23863  
tissue kallikrein (EC 3.4.21.35) S1, submaxillary - rat (fragment)  
N;Alternate names: glandular kallikrein; kininogenin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 22-Jul-1987 #sequence\_revision 22-Jul-1987 #text\_change 09-Jul-2004  
C;Accession: B23863  
R;Ashley, P.B.; Macdonald, R.J.  
Biochemistry 24, 4512-4520, 1985  
A;Title: Kallikrein-related mRNAs of the rat submaxillary gland: nucleotide sequences of  
A;Reference number: A23863; MUID:86051477; PMID:2998455  
A;Accession: B23863  
A;Molecule type: mRNA  
A;Residues: 1-156 <ASH>  
A;Cross-references: UNIPROT:P15950; GB:M11564; NID:G205031; PIDN:AAA41465.1; PID:G205032  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-148/Domain: trypsin homology (fragment) <TRY>  
  
Query Match 74.0%; Score 37; DB 2; Length 156;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 GPLVCRGTL 9  
		:
		:
		:
		:

Db 110 GPLICDGV 118

## RESULT 13

B23340  
tissue kallikrein (EC 3.4.21.35) 3 - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 22-Nov-1989 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: B23340  
R;Shai, S.Y.; Woodley-Miller, C.; Chao, J.; Chao, L.  
Biochemistry 28, 5334-5343, 1989  
A;Title: Characterization of genes encoding rat tonin and a kallikrein-like serine protease  
A;Reference number: A23340; MUID:89375248; PMID:2550051  
A;Accession: B23340  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-188 <SHA>  
A;Cross-references: UNIPROT:P15950; GB:M26534; NID:G206773; PIDN:AAA42080.1; PID:G206774;  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-180/Domain: trypsin homology (fragment) <TRY>

Query Match 74.0%; Score 37; DB 2; Length 188;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 142 GPLICDGV 150  
		:
		:
		:
		:

## RESULT 14

S45356  
probable serine proteinase (EC 3.4.21.-) ACO - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 14-Apr-1998 #sequence\_revision 08-May-1998 #text\_change 19-May-2000  
C;Accession: S45356  
R;Dihanach, M.; Spiess, M.  
Biochim. Biophys. Acta 1218, 225-228, 1994  
A;Title: A novel serine proteinase-like sequence from human brain.  
A;Reference number: S45356; MUID:94289486; PMID:8018728  
A;Accession: S45356  
A;Molecule type: mRNA  
A;Residues: 1-225 <DIH>  
A;Cross-references: EMBL:X75363; NID:9407137; PIDN:CAA53145.1; PID:G940540  
A;Experimental source: Alzheimer's disease patient brain cortex  
C;Genetics:  
A;Gene: ACO  
A;Introns: 175/3  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase

Query Match 74.0%; Score 37; DB 2; Length 225;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 180 GPLVCGGIL 188  
		:
		:
		:
		:

## RESULT 15

A27207  
tissue kallikrein (EC 3.4.21.35), prostatic - guinea pig  
N;Alternate names: glandular kallikrein  
C;Species: Cavia porcellus (guinea pig)  
C;Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 09-Jul-2004  
C;Accession: A27207  
R;Dunbar, J.C.; Bradshaw, R.A.  
Biochemistry 26, 3471-3478, 1987  
A;Title: Amino acid sequence of guinea pig prostate kallikrein.  
A;Reference number: A27207; MUID:88000549; PMID:3307909  
A;Accession: A27207

A;Molecule type: protein  
A;Residues: 1-239 <DUN>  
A;Cross-references: UNIPROT:P12323  
A;Note: 50-Trp was also found  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-231/Domain: trypsin homology <TRY>  
F;41.96.191/Active site: His, Asp, Ser #status predicted

Query Match 74.0%; Score 37; DB 2; Length 239;  
Best Local Similarity 66.7%; Pred.No. 21;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
|||:|  
Db 193 GPLICDGVL 201

Search completed: July 13, 2005, 17:31:31  
Job time : 13.5556 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-80  
Perfect score: 50  
Sequence: 1 GPLVCRGTL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	50	100.0	181	2	Q8NFV7	Q8nfv7 homo sapien
2	50	100.0	253	1	KLK7_HUMAN	P49862 homo sapien
3	50	100.0	258	2	Q8NSN9	Q8nsn9 homo sapien
4	47	94.0	258	2	Q63ZF4	Q63zf4 rattus norv
5	45	90.0	250	1	KLK9_HUMAN	Q9ukq9 homo sapien
6	45	90.0	251	2	Q9PBQ8	Q9dbq8 mus musculu
7	44	88.0	210	2	Q63Z11	Q63zil xenopus lae
8	44	88.0	251	1	KLKE_HUMAN	Q9p0g3 homo sapien
9	44	88.0	251	2	Q6B089	Q6b089 homo sapien
10	41	82.0	87	2	Q9CQ78	Q9cq78 m mus muscu
11	41	82.0	163	2	Q66H01	Q66h01 xenopus lae
12	41	82.0	249	2	Q9QYN4	Q9qyn4 mus musculu
13	41	82.0	250	2	Q63ZF2	Q63zf2 rattus norv
14	41	82.0	276	2	Q9QYN3	Q9qyn3 m hippostas
15	41	82.0	612	2	Q83WX7	Q83wx7 streptomyce
16	41	82.0	621	2	Q8G3L1	Q8g3l1 bifidobacte
17	41	82.0	3201	2	Q9F828	Q9f828 micromonosp
18	41	82.0	3546	2	Q9F830	Q9f830 micromonosp
19	40	80.0	180	2	Q7Q831	Q7q831 anopheles g
20	40	80.0	220	2	Q8NCW4	Q8ncw4 homo sapien
21	40	80.0	237	2	Q6GYJ5	Q6gyj5 struthio ca
22	40	80.0	254	2	Q9XSN6	Q9xen6 sus scrofa
23	40	80.0	257	2	Q6LDS3	Q6lds3 homo sapien
24	40	80.0	257	2	Q8QG86	Q8qg86 bothrops in
25	40	80.0	258	1	VF3_BOTJA	Q9pcu8 bothrops ja
26	40	80.0	260	1	NRPN_RAT	Q88780 rattus norv
27	40	80.0	260	2	Q71QH9	Q71qh9 trimeresuru
28	40	80.0	260	2	Q71Q14	Q71q14 trimeresuru
29	40	80.0	260	2	Q71QJ2	Q71qj2 trimeresuru
30	40	80.0	261	1	KLK2_HUMAN	P20151 homo sapien
31	40	80.0	261	1	KLK3_HUMAN	P07288 homo sapien

32	40	80.0	262	2	Q9XAC9	Q9xac9 streptomyce
33	40	80.0	420	2	Q7QZ35	Q7qz35 anopheles g
34	40	80.0	619	2	Q8KEH3	Q8keh3 chlorobium
35	39	78.0	155	2	Q9YBC2	Q9ybc2 aeropyrum p
36	39	78.0	247	2	Q9W7Q5	Q9w7q5 paralichthy
37	39	78.0	261	1	KLK1_RAT	P00758 rattus norv
38	39	78.0	265	2	Q68G17	Q68g17 rattus norv
39	39	78.0	297	2	Q9W174	Q9w174 drosophila
40	39	78.0	485	2	Q7PKK0	Q7pkk0 anopheles g
41	39	78.0	486	2	Q7FX74	Q7fx74 anopheles g
42	39	78.0	536	2	Q7PX72	Q7px72 anopheles g
43	38	76.0	119	2	Q9NR68	Q9nr68 homo sapien
44	38	76.0	228	2	Q6FHM3	Q6fhw3 homo sapien
45	38	76.0	234	2	Q9CV76	Q9cv76 mus musculu

#### ALIGNMENTS

##### RESULT 1

Q8NFV7 ID Q8NFV7 PRELIMINARY; PRT; 181 AA.  
AC Q8NFV7;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Kallikrein 7 short variant protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A.  
RC TISSUE=Ovarian carcinoma;  
RX MEDLINE=226223266; PubMed=12738725;  
RA Dong Y., Kaulahal A., Brattsand M., Nicklin J., Clements J.A.;  
RT "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer  
produces novel variants with potential as cancer biomarkers.";  
RL Clin. Cancer Res. 9:1710-1720(2003).  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AF411215; AA03663.1; -;  
DR HSSP; P00760; 1EZX.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 181 AA; 19887 MW; 86A28A03B80C2D78 CRC64;

Query Match 100.0%; Score 50; DB 2; Length 181;

Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9

Db 135 GPLVCRGTL 143

##### RESULT 2

KLK7\_HUMAN ID KLK7\_HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic  
 DE enzyme) (hSCCE).  
 GN Name=KLK7; Synonyms=PRSS6, SCCE;  
 OS Homo sapiens (Human)  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
 RC TISSUE=Skin;  
 RX MEDLINE=94308225; PubMed=8034709;  
 RA Hansson L., Stromqvist M., Backman A., Wallbrandt P., Carlstein A.,  
 RA Egelrud T.;  
 RT "Cloning, expression, and characterization of stratum corneum  
 RT chymotryptic enzyme. A skin-specific human serine proteinase.";  
 RL J. Biol. Chem. 269:19420-19426 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Keratinocytes;  
 RA Yousef G.M., Scorialas A., Diamandis E.P.;  
 RT "Molecular characterization, mapping and tissue expression of the  
 RT human stratum corneum chymotryptic enzyme gene.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepker B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130 (2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
 RA Wallbrandt P., Egelrud T.;  
 RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
 RT mice; a model for chronic itchy dermatitis.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95314630; PubMed=7794273;  
 RA Skjott A., Stromqvist M., Egelrud T.;  
 RT "Primary substrate specificity of recombinant human stratum corneum  
 RT chymotryptic enzyme.";  
 RL Biochem. Biophys. Res. Commun. 211:586-589 (1995).  
 CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive  
 CC structures in the cornified layer of the skin in the continuous  
 CC shedding of cells from the skin surface. Specific for amino acid  
 CC residues with aromatic side chains in the P1 position. SCCE  
 CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
 CC Phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
 CC activation of precursors to inflammatory cytokines.  
 CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
 CC expressed by keratinocytes in the epidermis. Very low levels are  
 CC also seen in the brain and kidney.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L33404; AAC37551.1; -;  
 DR EMBL; AF166330; AAD49718.1; -;  
 DR EMBL; AF243527; AAG33360.1; -;  
 DR EMBL; AF332583; AAK69624.1; -;  
 DR F01; A53968; A53968.  
 DR HSP; F00760; IEXX.  
 DR MEROPS; S01.300; -;

DR Genew; HGNC:6368; KLUK7.  
 DR H-InvDB; HIX0015373; -;  
 DR MIM; 604438; -;  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR GO; GO:0008544; P:epidermal differentiation; TAS.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Sec; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
 KW signal; Zymogen.  
 FT SIGNAL 1 22  
 FT PROPEP 23 29  
 FT CHAIN 30 253  
 FT ACT\_SITE 70 70  
 FT ACT\_SITE 112 112  
 FT ACT\_SITE 205 205  
 FT DISULFID 36 137  
 FT DISULFID 55 71  
 FT DISULFID ? 239  
 FT DISULFID 144 211  
 FT DISULFID 176 190  
 FT DISULFID 201 226  
 FT CARBOHYD 246 246  
 SQ SEQUENCE 253 AA; 27525 MW; 2D6B86B15A76A668 CRC64;  
 Query Match 100.0%; Score 50; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GPLVCRGTL 9  
 Db 207 GPLVCRGTL 215  
 RESULT 3  
 Q8N5N9 ID Q8N5N9 PRELIMINARY; PRT; 253 AA.  
 AC Q8N5N9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Stratum corneum chymotryptic enzyme, preproprotein.  
 GN Name=KLK7;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,

RT Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RA Strausberg R.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: Belongs to peptidase family S1.

DR EMBL; BC032005; AAH32005.1; -.

DR HSSP; P00760; 1E2X.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001254; Peptidase S1.

DR InterPro; IPR001314; Peptidase S1A.

DR InterPro; IPR009003; Pept\_Ser\_Cys.

DR Pfam; PF00089; Trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS0240; TRYPsin\_DOM; 1.

DR PROSITE; PS00134; TRYPsin\_HIS; UNKNOWN\_1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.

KW Hydrolase; Protease; Serine protease.

SQ SEQUENCE 253 AA; 27608 MW; 2D68BGA41B22A668 CRC64;

Query Match 100.0%; Score 50; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 207 GPLVCRGTL 215

RESULT 4

Q632F4

AC Q632F4 PRELIMINARY; PRT; 258 AA.

AD Q632F4;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Glandular kallikrein 9 precursor.

GN Name=Klk9;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=15203212; DOI=10.1016/j.ygeno.2004.01.009;

RA Olsson A.Y., Lilja H., Lundwall A.;

RT taxon-specific evolution of glandular kallikrein genes and

RT identification of a progenitor of prostate-specific antigen.";

RL Genomics 84:147-156(2004).

CC -!- MISCELLANEOUS: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ third party annotation (TPA) entry.

DR EMBL; BK001362; DAA05598.1; -.

KW Signal.

FT SIGNAL 1 19 Potential.

FT CHAIN 21 258 Glandular kallikrein 9.

SQ SEQUENCE 258 AA; 28581 MW; 5933AB83F6BE4461 CRC64;

Query Match 94.0%; Score 47; DB 2; Length 258;  
 Best Local Similarity 88.9%; Pred. No. 0.91;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 206 GPLVCRGTL 214

RESULT 5  
 KLK9\_HUMAN  
 ID KLK9\_HUMAN STANDARD; PRT; 250 AA.  
 AC Q9UKQ9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Kallikrein 9 precursor (EC 3.4.21.-) (Kallikrein-like protein 3) (KLK-L3).  
 DE L3).  
 GN Name=KLK9;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20118156; PubMed=10652563;  
 RA Yousef G.M., Luo L.-Y., Diamandis E.P.;

"Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4.";  
 RT Anticancer Res. 19:2843-2852(1999).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20247258; PubMed=10783266; DOI=10.1006/geno.2000.6159;  
 RA Yousef G.M., Diamandis E.P.;

"The expanded human kallikrein gene family: locus characterization and molecular cloning of a new member, KLK-L3.";  
 RT Genomics 65:184-194(2000).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=1054574; DOI=10.1016/S0378-1119(00)00382-6;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepker B., Wang K.;

"Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
 RT Gene 257:119-130(2000).  
 RL [4]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15057824; DOI=10.1038/nature02399;  
 RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J., Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M., Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E., Canepeel S., Carrano A.V., Caolite C., Chan Y.M., Christensen M., Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Dettler J.C., Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M., Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T., Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H., Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S., Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J., Nelson K., Nolan M., Ovcharenko I., Pitluck S., Pollard M., Popkie A.P., Predki P., Quan G., Ramirez L., Rash S., Retterer J., Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D., Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A., Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I., Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E., Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M., Rubin E.M., Lucas S.M.;

"The DNA sequence and biology of human chromosome 19.";  
 RL Nature 428:529-535(2004).  
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -!- TISSUE SPECIFICITY: Skin, thymus, trachea, cerebellum and spinal cord.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL; AF135026; AAD26427.2; --
DR EMBL; AF243527; AAG33362.1; --
DR EMBL; AC011473; AAG23255.1; --
DR HSSP; P00746; 1FDP.
DR MEROPS; S01.307; --
DR MIM; 605504; --
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0004252; E:serine-type endopeptidase activity; NAS.
DR GO; GO:0006508; F:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Serine protease; Signal.
FT SIGNAL 1 15 Potential.
FT CHAIN 16 250 Kallikrein 9.
FT ACT_SITE 63 63 Charge relay system (By similarity).
FT ACT_SITE 111 111 Charge relay system (By similarity).
FT ACT_SITE 204 204 Charge relay system (By similarity).
FT DISULFID 29 164 By similarity.
FT DISULFID 48 64 By similarity.
FT DISULFID 136 238 By similarity.
FT DISULFID 143 210 By similarity.
FT DISULFID 175 189 By similarity.
FT DISULFID 200 225 By similarity.
FT CARBOHYD 131 131 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 166 166 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 211 211 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 250 AA; 27512 MW; F27852458063E98B CRC64;

Query Match 90.0%; Score 45; DB 1; Length 250;
Best Local Similarity 88.9%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 206 GPLVCGNGL 214

RESULT 6
Q9DBQ8 PRELIMINARY; PRT; 251 AA.
AC Q9DBQ8;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,
DE clone:1200016C12 product:similar to KALLIKREIN 9 (EC 3.4.21.-)
DE (KALLIKREIN-LIKE PROTEIN 3) (KLK- L3).
GN Name=1200016C12Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;

```

```

RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai T.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AK004807; BAB23579.1; --
DR HSSP; P00746; 1FDP.
DR MEROPS; S01.307; --
DR MIM; MGI:1921082; 1200016C12Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_SIA.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 251 AA; 28153 MW; F4D667F8C80C4A23 CRC64;

Query Match 90.0%; Score 45; DB 2; Length 251;
Best Local Similarity 88.9%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 206 GPLVCGNGL 214

```

```

RESULT 7
Q63211 PRELIMINARY; PRT; 210 AA.
ID Q63211
DC 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082936; AAK82936.1; -
KW Hypothetical protein.
SQ SEQUENCE 210 AA; 22897 MW; 7543224C2227DA35 CRC64;

Query Match 88.0%; Score 44; DB 2; Length 210;
Best Local Similarity 88.9%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
DB 165 GPLVCRGEL 173

RESULT 8
KLKE_HUMAN
ID KLKE_HUMAN STANDARD; PRT; 251 AA.
AC Q9P0G3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Kallikrein 14 precursor (EC 3.4.21.-) (Kallikrein-like protein 6)

```

```

DE (KLK-L6).
CN Name=KLK14; Synonyms=KLKL6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yousef G.M., Diamandis E.P.;
RT "Molecular characterization, mapping, and tissue expression of KLKL6,
RT a hormonally regulated kallikrein-like gene.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=21250997; PubMed=11352573; DOI=10.1006/geno.2000.6490;
RA Hooper J.D., Bui L.T., Rae F.K., Harvey T.J., Myers S.A.,
RA Ashworth L.K., Clements J.A.;
RT "Identification and characterization of KLK14, a novel kallikrein
RT serine protease gene located on human chromosome 19q13.4 and expressed
RT in prostate and skeletal muscle.";
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=15057824; DOI=10.1038/nature02399;
RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,
RA Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M.,
RA Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,
RA Caenepeel S., Carrano A.V., Caciile C., Chan Y.M., Christensen M.,
RA Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Dettler J.C.,
RA Escobar J., Flowers D., Fotopulos D., Garcia C., Georgescu A.M.,
RA Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T.,
RA Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H.,
RA Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,
RA Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J.,
RA Nelson K., Nolan M., Ovcharenko I., Pittluck S., Pollard M.,
RA Popkie A.P., Predki P., Quan G., Ramirez L., Rash S., Retterer J.,
RA Rodriguez A., Rogers S., Salamon A., Salazar A., She X., Smith D.,
RA Slezak T., Soloviyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,
RA Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Buchhak I.,
RA Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E.,
RA Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,
RA Rubin E.M., Lucas S.M.;
RT "The DNA sequence and biology of human chromosome 19.";
RL Nature 428:529-535 (2004).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=20545474; PubMed=10969073; DOI=10.1074/jbc.M004525200;
RA Harvey T.J., Hooper J.D., Myers S.A., Stephenson S.A., Ashworth L.K.,
RA Clements J.A.;
RT "Tissue-specific expression patterns and fine mapping of the human
RT kallikrein (KLK) locus on proximal 19q13.4.";
RL J. Biol. Chem. 275:37397-37406 (2000).
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: High expression in brain, bone marrow and
CC fetal liver. Also expressed in liver, pancreas, fetal spleen,
CC prostate and skeletal muscle.
CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC
DR EMBL; AF161221; AAD50773.2; -
DR EMBL; AF283669; AAK48523.1; -
DR EMBL; AF283670; AAK48524.1; -
DR EMBL; AC011473; AAG23260.1; -
DR HSPB; P00760; 1EXZ.
DR MEROPS; S01.029; -

```

```

DR Genew; HGNC:6362; KLK14.
DR MIM; 606135; -.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0004252; P:serine-type endopeptidase activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HydroLase; Serine protease; Signal; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 24 Activation peptide (Potential).
FT CHAIN 25 251 Kallikrein 14.
FT ACT_SITE 67 67 Charge relay system (By similarity).
FT ACT_SITE 111 111 Charge relay system (By similarity).
FT ACT_SITE 204 204 Charge relay system (By similarity).
FT DISULFID 31 164 By similarity.
FT DISULFID 52 68 By similarity.
FT DISULFID 143 210 By similarity.
FT DISULFID 175 189 By similarity.
FT DISULFID 200 225 By similarity.
SQ SEQUENCE 251 AA; 27452 MW; 9087953BAFA7ED25 CRC64;

Query Match 88.0%; Score 44; DB 1; Length 251;
Best Local Similarity 88.9%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
DB 206 GPLVCRGQL 214

RESULT 9
QB089 PRELIMINARY; PRT; 251 AA.
AC Q6B089
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Kallikrein 14, preproprotein.
GN Name=KLK14;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Director MGC Project;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; BC074905; AAH74905.1; -.
DR EMBL; BC074904; AAH74904.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW HydroLase; Protease; Serine protease.
SQ SEQUENCE 251 AA; 27507 MW; C353A7DB81960BD2 CRC64;

Query Match 88.0%; Score 44; DB 2; Length 251;
Best Local Similarity 88.9%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
DB 206 GPLVCRGQL 214

RESULT 10
Q9CQ78 PRELIMINARY; PRT; 87 AA.
AC Q9CQ78;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DE library, clone:2310081E03 product:protease, serine, 20, full insert
DE sequence (Mus musculus adult male tongue cDNA, RIKEN full-length
DE enriched library, clone:2310021N04 product:protease, serine, 20, full
DE insert sequence) (Fragment).
GN Name=2310015108Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
```



RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RL prepare full-length cDNA libraries for rapid discovery of new genes.";  
RN Genome Res. 10:1617-1630(2000).  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=20530913; PubMed=1076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto K., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito K., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito K., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [8]  
DR EMBL; AK019102; BAB31548.1; -;  
DR EMBL; AK009451; BAB26297.1; -;  
DR HSSP; P00760; 1EXZ.  
DR MEROPS; S01.257; -;  
DR MGD; MGI:192977; 2310015108Rik.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON TER 1  
SQ SEQUENCE 87 AA; 9549 MW; B988D0C62926EAA CRC64;

Query Match 82.0%; Score 41; DB 2; Length 87;  
Best Local Similarity 77.8%; Pred. No. 4.3;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9

Db ||||| |:-|  
42 GPLVNGSL 50  
RESULT 11  
Q66H01 PRELIMINARY;  
AC Q66H01 PRT; 163 AA.  
DT 25-OCT-2004 (TRENBLrel. 28, Created)  
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
DE MGC97901 protein.  
GN Name=MGC97901;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RL Dev. Dyn. 225:384-391(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Gerhard D.S.;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; BC082114; AAH82114.1; -;  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 163 AA; 17881 MW; DC90CAEAA755BAFB CRC64;

Query Match 82.0%; Score 41; DB 2; Length 163;  
Best Local Similarity 77.8%; Pred. No. 8.1;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 GPLVCRGTL 9
Db 118 GPLVCNGSL 126

RESULT 12
Q9QYN4 PRELIMINARY; PRT; 249 AA.
AC Q9QYN4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Hippostasin.
GN Name=2310015108Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
EX MEDLINE=20525460; PubMed=11072088; DOI=10.1016/S0167-4781(00)00206-2;
RA Mitsui S., Okui A., Kominami K., Uemura H., Yamaguchi N.;
RT "cDNA cloning and tissue-specific splicing variants of mouse
RL Biochim. Biophys. Acta 1494:206-210(2000).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AB016226; BAA88825.1; -.
DR HSSP; P00760; 1EZX.
DR MEROPS; S01.257; -.
DR MGD; MGI:1929977; 2310015108Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00889; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 249 AA; 27604 MW; F9FF9CB457D727D5 CRC64;

Query Match 82.0%; Score 41; DB 2; Length 249;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
Db 204 GPLVCNGSL 212

RESULT 13
Q632F2 PRELIMINARY; PRT; 250 AA.
AC Q632F2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Glandular kallikrein 11 precursor.
GN Name=K1k11;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX PubMed=15203212; DOI=10.1016/j.ygeno.2004.01.009;
RA Olsson A.Y., Lilja H., Lundwall A.;
RT "Taxon-specific evolution of glandular kallikrein genes and
RL identification of a progenitor of prostate-specific antigen.";
RL Genomics 84:147-156(2004).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK001364; DAA05600.1; -.
KW Signal.
FT SIGNAL 1 18 Potential.
FT CHAIN 22 250 Glandular kallikrein 11.
SQ SEQUENCE 250 AA; 27630 MW; 247B29D3F0DF8F1 CRC64;

Query Match 82.0%; Score 41; DB 2; Length 250;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
Db 205 GPLVCNGSL 213

RESULT 14
Q9QYN3 PRELIMINARY; PRT; 276 AA.
AC Q9QYN3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hippostasin prostate type (Mus musculus adult male tongue cDNA, RIKEN
DE full-length enriched library, clone:2310015108 product:protease,
DE serine, 20, full insert sequence) (Mus musculus adult male tongue
DE cDNA, RIKEN full-length enriched library, clone:2310040F07
DE product:protease, serine, 20, full insert sequence).
GN Name=2310015108Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=20525460; PubMed=11072088; DOI=10.1016/S0167-4781(00)00206-2;
RA Mitsui S., Okui A., Kominami K., Uemura H., Yamaguchi N.;
RT "cDNA cloning and tissue-specific splicing variants of mouse
RL Biochim. Biophys. Acta 1494:206-210(2000).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
```

```

RL Nature 420:563-573(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuul T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hitamoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AB016227; BAA36955.1; -
DR EMBL; AK009360; BAB26241.2; -
DR EMBL; AK009720; BAB26461.2; -
DR HSSP; P00760; 1EXZ.
DR MEROPS; S01.257; -
DR MGD; MGI:1929977; 2310015I08R1k.
GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Protease; Serine protease.
SQ SEQUENCE 276 AA; 30753 MW; 90BDC03A8AB178D6 CRC64;

Query Match 82.0%; Score 41; DB 2; Length 276;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 231 GPLVCGSL 239

RESULT 15
Q83WX7 PRELIMINARY; PRT; 612 AA.
AC Q83WX7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Streptomyces rochei (Streptomyces parvullus).
OG Plasmid pSLA2-L.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7434ANA;
RX MEDLINE=22676866; PubMed=12791134;
RA Kinashi H.; Hiratsu K., Suwa M., Ishii T., Sugino F., Yamada K.,
RT "The large linear plasmid pSLA2-L of Streptomyces rochei has an
RT unusually condensed gene organization for secondary metabolism.";
RL Mol. Microbiol. 48:1501-1510(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7434ANA;
RX MEDLINE=20408175; PubMed=10954087;
RA Hiratsu K., Mochizuki S., Kinashi H.;
RT "Cloning and analysis of the replication origin and the telomeres of
RT the large linear plasmid pSLA2-L in Streptomyces rochei.";
RL Mol. Gen. Genet. 263:1015-1021(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=7434ANA;
RX MEDLINE=20231737; PubMed=10767533; DOI=10.1016/S0378-1119(00)00060-3;
RA Suwa M., Sugino H., Sasaoka A., Mori E., Fujii S., Shinkawa H.,
RA Nini O., Kinashi H.;
RT "Identification of two polyketide synthase gene clusters on the linear
RT plasmid pSLA2-L in Streptomyces rochei.";
RL Gene 246:123-131(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=7434ANA;
RX MEDLINE=99053144; PubMed=9836424;
RA Kinashi H., Fujii S., Hatani A., Kurokawa T., Shinkawa H.;
RT "Physical mapping of the linear plasmid pSLA2-L and localization of
RT the eryAI and actI homologs.";
RL Biosci. Biotechnol. Biochem. 62:1892-1897(1998).
DR EMBL; AB088224; BAC76588.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 612 AA; 66525 MW; 98F0B1EB9D3110FC CRC64;

Query Match 82.0%; Score 41; DB 2; Length 612;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 250 GPLVCAGTV 258

Search completed: July 13, 2005, 17:29:27
Job time : 64.2222 secs

```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-80

Perfect score: 50

Sequence: 1 GPLVCRGTL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	9	3	US-09-502-600-80
2	50	100.0	9	4	US-09-918-243-80
3	50	100.0	224	3	US-08-944-483-33
4	50	100.0	225	2	US-08-557-146-12
5	50	100.0	225	2	US-09-027-337-4
6	50	100.0	225	2	US-09-154-344-12
7	50	100.0	225	4	US-09-644-600-4
8	50	100.0	225	4	US-09-654-600A-4
9	50	100.0	253	2	US-08-557-146-2
10	50	100.0	253	2	US-08-824-874-3
11	50	100.0	253	2	US-09-154-344-2
12	50	100.0	253	3	US-08-930-188-2
13	50	100.0	253	3	US-09-210-084-3
14	50	100.0	253	4	US-09-764-762-3
15	50	100.0	253	5	PCT-US96-04294-2
16	50	100.0	265	4	US-09-949-016-7716
17	41	82.0	3201	4	US-09-679-279-15
18	41	82.0	3546	4	US-09-679-279-13
19	40	80.0	20	1	US-08-473-228A-19
20	40	80.0	20	3	US-09-146-831-19
21	40	80.0	20	5	PCT-US96-09303-19
22	40	80.0	237	1	US-08-096-946-10
23	40	80.0	237	1	US-08-096-946-11
24	40	80.0	237	2	US-08-844-024-2
25	40	80.0	237	2	US-08-718-547-2
26	40	80.0	237	3	US-08-768-859A-1
27	40	80.0	237	3	US-08-768-859A-16

28 40 80.0 237 3 US-08-768-859A-21 Sequence 21, Appli

29 40 80.0 237 3 US-08-767-820A-16 Sequence 1, Appli

30 40 80.0 237 3 US-08-767-820A-16 Sequence 16, Appli

31 40 80.0 237 3 US-08-767-820A-21 Sequence 21, Appli

32 40 80.0 237 3 US-08-622-046B-1 Sequence 1, Appli

33 40 80.0 237 3 US-08-622-046B-7 Sequence 7, Appli

34 40 80.0 237 3 US-08-622-046B-12 Sequence 12, Appli

35 40 80.0 237 3 US-08-944-483-37 Sequence 37, Appli

36 40 80.0 237 3 US-08-944-483-38 Sequence 38, Appli

37 40 80.0 237 3 US-09-100-264-1 Sequence 1, Appli

38 40 80.0 237 3 US-09-100-264-3 Sequence 3, Appli

39 40 80.0 237 3 US-09-100-264-12 Sequence 12, Appli

40 40 80.0 237 4 US-09-303-339-2 Sequence 2, Appli

41 40 80.0 237 4 US-08-843-076D-1 Sequence 1, Appli

42 40 80.0 237 4 US-08-843-076D-7 Sequence 7, Appli

43 40 80.0 237 4 US-08-843-076D-8 Sequence 8, Appli

44 40 80.0 237 4 US-09-303-208-1 Sequence 1, Appli

45 40 80.0 237 5 PCT-US94-07329-10 Sequence 10, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-502-600-80

; Sequence 80, Application US/09502600A

; Patent No. 6294344

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of

; TITLE OF INVENTION: Ovarian Cancer

; FILE REFERENCE: D6223CIP-C

; CURRENT FILING DATE: 2000-02-11

; CURRENT APPLICATION NUMBER: US/09/502,600A

; PRIOR APPLICATION NUMBER: 09/039,211

; PRIOR FILING DATE: 03-14-1998

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 80

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Residues 207-215 of the SCCE protein

US-09-502-600-80

Query Match Similarity 100.0%; Score 50; DB 3; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.1e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9

Db 1 GPLVCRGTL 9

##### RESULT 2

US-09-918-243-80

; Sequence 80, Application US/09918243

; Patent No. 6627403

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; APPLICANT: Cannon, Martin J.

; APPLICANT: Santin, Alessandro

; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP

; CURRENT APPLICATION NUMBER: US/09/918,243

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US

; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 80

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 207-215 of the SCCE protein  
US-09-918-243-80

Query Match 100.0%; Score 50; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 1 GPLVCRGTL 9

RESULT 3  
US-08-944-483-33  
; Sequence 33, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; TITLE OF INVENTION: OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 224 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6232456e  
US-08-944-483-33

Query Match 100.0%; Score 50; DB 3; Length 224;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 1 GPLVCRGTL 186

US-08-557-146-12  
; Sequence 12, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sterner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 225 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
US-08-557-146-12

Query Match 100.0%; Score 50; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 179 GPLVCRGTL 187

RESULT 5  
US-09-027-337-4  
; Sequence 4, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotoashi  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in  
; TITLE OF INVENTION: Breast and Ovarian Carcinomas  
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027,337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to

```
; OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4
Query Match      100.0%; Score 50; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
Db 179 GPLVCRGTL 187

RESULT 6
US-09-154-344-12
; Sequence 12, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-09-154-344-12
Query Match      100.0%; Score 50; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
Db 179 GPLVCRGTL 187

RESULT 7
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.

; OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4
Query Match      100.0%; Score 50; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
Db 179 GPLVCRGTL 187

RESULT 8
US-09-654-600A-4
; Sequence 4, Application US/09654600A
; Patent No. 6649741
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/654,600A
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
; 09/027,337
; PRIOR FILING DATE: 1999-10-20
; 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-654-600A-4
Query Match      100.0%; Score 50; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPLVCRGTL 9
Db 179 GPLVCRGTL 187

RESULT 9
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-557-146-2
Query Match 100.0%; Score 50; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215

RESULT 10
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
```

```
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
;
US-08-824-874-3
Query Match 100.0%; Score 50; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215

RESULT 11
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-154-344-2
Query Match 100.0%; Score 50; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215
```



RESULT 12  
US-08-930-188-2  
; Sequence 2, Application US/08930188  
; Patent No. 6093397  
; GENERAL INFORMATION:  
; APPLICANT: Dixon, Eric P.  
; APPLICANT: Johnstone, Edward M.  
; APPLICANT: Little, Shella P.  
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
; RELATED NUCLEIC ACIDS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,188  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/416,257  
; FILING DATE: 04-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blalock, Donna K.  
; REGISTRATION NUMBER: 38,082  
; REFERENCE/DOCKET NUMBER: X9239  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-277-1090  
; TELEFAX: 317-276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-930-188-2  
Query Match 100.0%; Score 50; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GPLVCRGTL 9  
Db 207 GPLVCRGTL 215  
RESULT 13  
US-09-210-084-3  
; Sequence 3, Application US/09210084  
; Patent No. 6197511  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,084  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 532504  
US-09-210-084-3  
Query Match 100.0%; Score 50; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GPLVCRGTL 9  
Db 207 GPLVCRGTL 215  
RESULT 14  
US-09-764-762-3  
; Sequence 3, Application US/09764762  
; Patent No. 6472195  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/764,762  
; FILING DATE: 16-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/210,084  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 532504  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-764-762-3

Query Match 100.0%; Score 50; DB 4; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 207 GPLVCRGTL 215  
|||||

RESULT 15  
PCT-US96-04294-2  
; Sequence 2, Application PC/TUS9604294  
; GENERAL INFORMATION:  
; APPLICANT: Dixon, Eric P.  
; APPLICANT: Johnstone, Edward M.  
; APPLICANT: Little, Sheila P.  
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
; TITLE OF INVENTION: RELATED NUCLEIC ACIDS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/04294  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/416,257  
; FILING DATE: 04-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blalock, Donna K.  
; REGISTRATION NUMBER: 38,082  
; REFERENCE/DOCKET NUMBER: X9239  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-277-1090  
; TELEFAX: 317-276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US96-04294-2

Query Match 100.0%; Score 50; DB 5; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
c |||||

Db 207 GPLVCRGTL 215

Search completed: July 13, 2005, 17:34:28  
Job time : 19.4444 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds

(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-80

Perfect score: 50

Sequence: 1 GPLVCRGTL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1756220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	50	100.0	9	9	US-09-918-243-80
2	50	100.0	9	9	US-09-905-083-80
3	50	100.0	9	15	US-10-372-521-80
4	50	100.0	9	15	US-10-831-075-80
5	50	100.0	181	15	US-10-262-511-98
6	50	100.0	198	15	US-10-262-511-96
7	50	100.0	224	11	US-09-789-210-33
8	50	100.0	224	15	US-10-262-511-104
9	50	100.0	225	15	US-10-600-187-4
10	50	100.0	247	15	US-10-262-511-102
11	50	100.0	250	15	US-10-262-511-92
					Sequence 80, Appl
					Sequence 80, Appl
					Sequence 80, Appl
					Sequence 96, Appl
					Sequence 98, Appl
					Sequence 33, Appl
					Sequence 104, Appl
					Sequence 4, Appli
					Sequence 102, Appl
					Sequence 92, Appl

12	50	100.0	252	15	US-10-262-511-94	Sequence 94, Appl
13	50	100.0	253	9	US-09-888-615-98	Sequence 98, Appl
14	50	100.0	253	9	US-09-764-762-3	Sequence 3, Appli
15	50	100.0	253	14	US-10-071-214-2	Sequence 2, Appli
16	50	100.0	253	14	US-10-071-214-48	Sequence 48, Appl
17	50	100.0	253	14	US-10-264-283-90	Sequence 90, Appl
18	50	100.0	253	15	US-10-295-027-498	Sequence 498, App
19	50	100.0	253	15	US-10-173-999-48	Sequence 48, Appl
20	50	100.0	253	16	US-10-408-765A-639	Sequence 639, App
21	50	100.0	253	16	US-10-643-795A-95	Sequence 95, Appl
22	50	100.0	253	17	US-10-948-518-95	Sequence 95, Appl
23	50	100.0	253	17	US-10-868-490A-1	Sequence 1, Appli
24	50	100.0	257	15	US-10-344-394-38	Sequence 38, Appl
25	47	94.0	249	14	US-10-071-214-47	Sequence 47, Appl
26	45	90.0	247	9	US-09-764-898-284	Sequence 284, App
27	45	90.0	247	9	US-09-764-847-742	Sequence 742, App
28	45	90.0	247	14	US-10-092-154-742	Sequence 742, App
29	45	90.0	250	15	US-10-055-569A-69	Sequence 69, Appl
30	45	90.0	250	15	US-10-344-394-29	Sequence 29, Appl
31	45	90.0	251	9	US-09-764-898-212	Sequence 212, App
32	45	90.0	259	15	US-10-072-012-222	Sequence 222, App
33	44	88.0	243	14	US-10-071-214-46	Sequence 46, Appl
34	44	88.0	251	15	US-10-344-394-27	Sequence 27, Appl
35	44	88.0	267	17	US-10-480-988-20	Sequence 20, Appl
36	42	84.0	89	15	US-10-424-599-210636	Sequence 210636,
37	41	82.0	23	14	US-10-071-214-6	Sequence 6, Appli
38	41	82.0	80	16	US-10-425-115-348061	Sequence 348061,
39	41	82.0	249	15	US-10-055-569A-68	Sequence 68, Appl
40	41	82.0	276	15	US-10-055-569A-67	Sequence 67, Appl
41	40	80.0	54	10	US-09-860-739-5	Sequence 5, Appli
42	40	80.0	82	16	US-10-437-963-182803	Sequence 182803,
43	40	80.0	125	14	US-10-106-698-5946	Sequence 5946, Ap
44	40	80.0	218	15	US-10-115-479-12	Sequence 12, Appl
45	40	80.0	237	10	US-09-860-739-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-918-243-80  
; Sequence 80, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 80  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 207-215 of the SCCE protein  
US-09-918-243-80

Query Match 100.0%; Score 50; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GPLVCRGTL 9  
Db 1 GPLVCRGTL 9

RESULT 2

```
US-09-905-083-80
; Sequence 80, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 80
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 207-215 of the SCCE protein
US-09-905-083-80

Query Match      100.0%; Score 50; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 1 GPLVCRGTL 9

RESULT 3
US-10-372-521-80
; Sequence 80, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 80
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 207-215 of the SCCE protein
US-10-372-521-80

Query Match      100.0%; Score 50; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 1 GPLVCRGTL 9

RESULT 4
US-10-831-075-80
; Sequence 80, Application US/10831075
; Publication No. US20040224891A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
```

```
; FILE REFERENCE: D6223CIP/C/D/CIP3
; CURRENT APPLICATION NUMBER: US/10/831,075
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 10/372,521
; PRIOR FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 80
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 207-215 of the SCCE protein
US-10-831-075-80

Query Match      100.0%; Score 50; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 1 GPLVCRGTL 9

RESULT 5
US-10-262-511-98
; Sequence 98, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
```

; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 98  
; LENGTH: 181  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-98

Query Match 100.0%; Score 50; DB 15; Length 181;  
Best Local Similarity 100.0%; Pred. No. 0.97; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1 GPLVCRGTL 9  
Db 133 GPLVCRGTL 141

RESULT 6  
US-10-262-511-96  
; Sequence 96, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spyttek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 96  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-96

Query Match 100.0%; Score 50; DB 15; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.1; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1 GPLVCRGTL 9  
Db 152 GPLVCRGTL 160

RESULT 7  
US-09-789-210-33  
; Sequence 33, Application US/09789210  
; Publication No. US20040241646A1  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; COLPITTS, TRACEY L.  
; FRIEDMAN, PAULA N.  
; GRANADOS, EDWARD N.  
; KLASS, MICHAEL R.  
; RUSSELL, JOHN C.  
; STEWART, KENT D.  
; STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/789,210  
; FILING DATE: 20-Feb-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/944,483  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623

TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 224 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: None  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-09-789-210-33

Query Match 100.0%; Score 50; DB 11; Length 224;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 178 GPLVCRGTL 186

## RESULT 8

US-10-262-511-104  
; Sequence 104, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; PRIOR FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260

TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 224 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: None  
SEQUENCE DESCRIPTION: SEQ ID NOS: 439  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 104  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-104

Query Match 100.0%; Score 50; DB 15; Length 224;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 193 GPLVCRGTL 201

## RESULT 9

US-10-600-187-4  
; Sequence 4, Application US/10600187  
; Publication No. US20040086910A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotoshi  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
; FILE REFERENCE: D6064CIP/D  
; CURRENT APPLICATION NUMBER: US/10/600,187  
; CURRENT FILING DATE: 2003-06-20  
; PRIOR APPLICATION NUMBER: US/09/654,600A  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 09/421,213  
; 09/027,337  
; PRIOR FILING DATE: 1999-10-20  
; 1998-02-20  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: SCCE  
US-10-600-187-4

Query Match 100.0%; Score 50; DB 15; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
Db 179 GPLVCRGTL 187

## RESULT 10

US-10-262-511-102  
; Sequence 102, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Ort, Tatiana  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Zernhusen, Bryan D.  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Pena, Carol E. A.  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Shimkets, Richard A.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Berghs, Constance  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-462C  
 ; CURRENT APPLICATION NUMBER: US/10/262,511  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/373,815  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,642  
 ; PRIOR FILING DATE: 2002-05-17  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,038  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 102  
 ; LENGTH: 247  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-102

Query Match 100.0%; Score 50; DB 15; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
 Db 193 GPLVCRGTL 201

RESULT 11  
 US-10-262-511-92  
 ; Sequence 92, Application US/10262511  
 ; Publication No. US2004003823A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Ju, Jingfang  
 ; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Saaha)  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Ort, Tatiana  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Zernhusen, Bryan D.  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Pena, Carol E. A.  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Shimkets, Richard A.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Berghs, Constance  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-462C  
 ; CURRENT APPLICATION NUMBER: US/10/262,511  
 ; CURRENT FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/373,815  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,642  
 ; PRIOR FILING DATE: 2002-05-17  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,038  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/373,260  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO 92  
 ; LENGTH: 250  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-92

Query Match 100.0%; Score 50; DB 15; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9  
 Db 204 GPLVCRGTL 212

RESULT 12  
 US-10-262-511-94  
 ; Sequence 94, Application US/10262511  
 ; Publication No. US2004003823A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.

```
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 94
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-94

Query Match 100.0%; Score 50; DB 15; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 206 GPLVCRGTL 214

RESULT 13
US-09-888-615-98
; Sequence 98, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Query Match 100.0%; Score 50; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215

RESULT 14
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
```

```
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 94
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-94

Query Match 100.0%; Score 50; DB 15; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 206 GPLVCRGTL 214

RESULT 13
US-09-888-615-98
; Sequence 98, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Query Match 100.0%; Score 50; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215

RESULT 14
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
```



```
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3
Query Match      100.0%; Score 50; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215

RESULT 15
US-10-071-214-2
; Sequence 2, Application US/10071214
; Publication No. US2003006099A1
; GENERAL INFORMATION:
; APPLICANT: HANSSON, Lennart
; APPLICANT: EGELURUD, Torbjorn
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN
; FILE REFERENCE: HANSSON=3A
; CURRENT APPLICATION NUMBER: US/10/071,214
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,422
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: DK PA 2001 00218
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-071-214-2

Query Match      100.0%; Score 50; DB 14; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPLVCRGTL 9
Db 207 GPLVCRGTL 215

Search completed: July 13, 2005, 18:23:41
Job time : 65.8889 secs
```

This page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083A-86

Perfect score: 41

Sequence: 1 MARSLLLPL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

1: geneseqp1980s.\*

2: geneseqp1990s.\*

3: geneseqp2000s.\*

4: geneseqp2001s.\*

5: geneseqp2002s.\*

6: geneseqp2003as.\*

7: geneseqp2003bs.\*

8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	AAE08291	Aae08291 Human str
2	41	100.0	9	ADR68847	Adr68847 Human str
3	41	100.0	136	ABG23378	Abg23378 Novel hum
4	41	100.0	198	ADA05736	Ada05736 Human NOV
5	41	100.0	198	ADN62900	Adn62900 Human NOV
6	41	100.0	253	AAR67888	Aar67888 Human str
7	41	100.0	253	AAW05383	Aaw05383 Human amy
8	41	100.0	253	ABB84421	Abb84421 Human SCC
9	41	100.0	253	ABB84406	Abb84406 Human SCC
10	41	100.0	253	AAU82740	Aau82740 Amino aci
11	41	100.0	253	ABU07440	Abu07440 Protein d
12	41	100.0	253	ABU07471	Abu07471 Protein d
13	41	100.0	253	ABR58471	AbR58471 Human str
14	41	100.0	253	ADB80484	AdB80484 Ovarian c
15	41	100.0	253	ADJ68833	AdJ68833 Human hea
16	41	100.0	253	ADN39180	Adn39180 Cancer/an
17	41	100.0	253	ADL06515	AdL06515 Human tum
18	41	100.0	253	ADN04182	Adn04182 Antipsori
19	41	100.0	253	ADR72880	Adr72880 Human ova
20	41	100.0	257	AAE21326	Aae21326 Human HSC
21	37	90.2	752	ADI28864	Adi28864 Human ser
22	36	87.8	9	AAE08320	Aae08320 Human str
23	36	87.8	9	ADR68877	Adr68877 Human str
24	36	87.8	818	ABB98135	Abb98135 Human PWM
25	36	87.8	818	AAU82753	Aau82753 Amino aci

26	35	85.4	475	4	AAE10657	Aae10657 Secreted
27	35	85.4	475	4	AAE02609	Aae02609 Human sec
28	35	85.4	475	5	ABB78618	Abb78618 Secreted
29	35	85.4	518	2	AAW61362	Aaw61362 Aspartic
30	35	85.4	518	2	AAV13799	Aay13799 Human asp
31	35	85.4	518	2	AAV22239	Aay22239 Human CSP
32	35	85.4	518	2	AAV41714	Aay41714 Human PRO
33	35	85.4	518	3	AAV88424	Aay88424 Human asp
34	35	85.4	518	3	AAW44270	Aaw44270 Human PRO
35	35	85.4	518	4	AAU07201	Aau07201 Human asp
36	35	85.4	518	4	AAE10628	Aae10628 Human asp
37	35	85.4	518	4	AAE10656	Aae10656 Human-Asp
38	35	85.4	518	4	AAE06858	Aae06858 Human asp
39	35	85.4	518	4	AAE02608	Aae02608 Human asp
40	35	85.4	518	4	AAE02580	Aae02580 Human asp
41	35	85.4	518	4	AAU29059	Aau29059 Human PRO
42	35	85.4	518	4	AAU06602	Aau06602 Human Asp
43	35	85.4	518	5	ABB06531	Abb06531 Human asp
44	35	85.4	518	5	ABB78589	Abb78589 Human Asp
45	35	85.4	518	5	ABB78617	Abb78617 Human Asp

## ALIGNMENTS

RESULT 1  
AAE08291  
ID AAE08291 standard; peptide; 9 AA.  
XX  
AC AAE08291;  
XX  
DT 01-NOV-2001 (first entry)  
XX  
DE Human stratum corneum chymotrypsin enzyme peptide #56 (residues 1-9).  
XX  
KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;  
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;  
KW antisense therapy; malignant hyperplasia.  
XX  
OS Homo sapiens.  
XX  
PN WO200159158-A1.  
XX  
PD 16-AUG-2001.  
XX  
PF 07-FEB-2001; 2001WO-US003977.  
XX  
PR 11-FEB-2000; 2000US-00502600.  
XX  
PA (UYAR-) UNIV ARKANSAS.  
XX  
PI O'brien TJ;  
XX  
DR WPI; 2001-514676/56.  
XX  
PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.  
XX  
PS Claim 25; Page 114; 127pp; English.  
XX  
CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer.  
XX  
CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is human SCCE peptide

SQ Sequence 9 AA;

Query Match 100.0%; Score 41; DB 4; Length 9;

```
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
DB 1 MARSLLLPL 9

RESULT 2
ADR68847
ID ADR68847 standard; peptide; 9 AA.
AC ADR68847;
XX
DT 02-DEC-2004 (first entry)
DE Human stratum corneum chymotrytic enzyme peptide fragment SEQ ID NO:86.
XX
KW serine protease; stratum corneum chymotrytic enzyme; SCCE;
KW immune response; ovarian cancer; lung cancer; prostate cancer;
KW pancreatic cancer; colon cancer.
XX
OS Homo sapiens.
XX
FN WO2004075723-A2.
XX
PD 10-SEP-2004.
XX
PF 20-FEB-2004; 2004WO-US005134.
XX
PR 21-FEB-2003; 2003US-00372521.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI O'brien TJ, Cannon MJ, Santin A;
XX
DR WPI; 2004-653294/63.
XX
PT Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for
PT vaccinating an individual against SCCE, and in monitoring and developing
PT immunotherapies for ovarian and other malignancies.
XX
PS Claim 5; SEQ ID NO 86; 117pp; English.
XX
CC The invention relates to a novel method for vaccinating an individual
CC against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating
CC an individual with a SCCE peptide, which elicits an immune response in
CC the individual. A peptide of the invention acts as a stratum corneum
CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating
CC an individual against SCCE, particularly an individual having, suspected
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon
CC cancer. The oligonucleotide is useful for treating a neoplastic state in
CC an individual, such as ovarian, breast, lung, colon, prostate, or
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The
CC peptides are also useful in the monitoring and development of
CC immunotherapies for ovarian and other malignancies. The present sequence
CC represents a peptide fragment of serine protease SCCE (stratum corneum
CC chymotrytic enzyme).
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 41; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
DB 1 MARSLLLPL 9

RESULT 3
ABG23378
ID ABG23378 standard; protein; 136 AA.
XX
AC
```

---

```
XX
AC ABG23378;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #23369.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS87565.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 53737; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 136 AA;
Query Match 100.0%; Score 41; DB 4; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
DB 7 MARSLLLPL 15

RESULT 4
ADA05736
ID ADA05736 standard; protein; 198 AA.
XX
AC ADA05736;
```

XX DT 06-NOV-2003 (first entry)

XX DE Human NOV18c protein SEQ ID NO:96.

XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

XX KW immunomodulator; cytostatic; neurotropic; neuroprotective;

XX KW antiparkinsonian; antilipaeamic; gene therapy; human disease;

XX KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

XX KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX OS Homo sapiens.

XX PN WO2003029424-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031373.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 09-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

XX PR 19-APR-2002; 2002US-0373826P.

XX PR 19-APR-2002; 2002US-0373884P.

XX PR 22-APR-2002; 2002US-0374977P.

XX PR 16-MAY-2002; 2002US-0381037P.

XX PR 16-MAY-2002; 2002US-0381038P.

XX PR 16-MAY-2002; 2002US-0381042P.

XX PR 17-MAY-2002; 2002US-0381642P.

XX PR 28-MAY-2002; 2002US-0383656P.

XX PR 29-MAY-2002; 2002US-0383831P.

XX PR 25-JUN-2002; 2002US-0391335P.

XX PR 01-OCT-2002; 2002US-00262511.

XX PA (CURA-) CURAGEN CORP.

XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

XX PI Patturajan M, Sytek KA, Edinger SR, Ellerman K, Malyankar UM;

XX PI Ott T, Gorman L, Zerkusen BD, Anderson DW, Zhong M, Catterton E;

XX PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;

XX PI Shinkets RA, Rotherberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

XX PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX DR WPI; 2003-381626/36.

XX DR N-P5DB; ADA05735.

XX PT New NOVX polypeptides and nucleic acids, useful for diagnosing,

XX PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

XX PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

XX PT pharmacogenomics.

XX PS Claim 1; Page 170; 586pp; English.

XX CC The present invention describes NOVX proteins, where X can be 1 to 55

XX CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

XX CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described

CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or

CC aberrant physiological interactions of the polypeptide; (11) a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating

CC the activity of the polypeptide described above; (13) methods of treating

CC or preventing a pathology associated with the above polypeptide in a

CC mammal; and (14) a method for producing the above polypeptide. NOVX

CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian

CC and antilipaeamic activities, and can be used in gene therapy. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic

CC disorders such as diabetes or obesity, infections, cachexia, cancer,

CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation

CC probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence represents a human NOVX from the

CC present invention.

XX SQ Sequence 198 AA;

Query Match 100.0%; Score 41; DB 6; Length 198;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9

Db 1 MARSLLLPL 9

RESULT 5

ID ADN62900 standard; protein; 198 AA.

XX AC ADN62900;

XX DT 01-JUL-2004 (first entry)

XX DE Human NOV18c.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

XX KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;

XX KW Alzheimer's disease; Parkinson's disease; immune disorder;

XX KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;

XX KW wasting disorder.

XX OS Homo sapiens.

XX PN US2004038223-A1.

XX PD 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 05-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

09-OCT-2001; 2001US-0328056P.  
 12-OCT-2001; 2001US-0328849P.  
 15-OCT-2001; 2001US-0329414P.  
 17-OCT-2001; 2001US-0330142P.  
 18-OCT-2001; 2001US-0330309P.  
 22-OCT-2001; 2001US-0341058P.  
 24-OCT-2001; 2001US-0339266P.  
 24-OCT-2001; 2001US-0343629P.  
 29-OCT-2001; 2001US-0349575P.  
 01-NOV-2001; 2001US-0346357P.  
 17-APR-2002; 2002US-0373260P.  
 19-APR-2002; 2002US-0373815P.  
 19-APR-2002; 2002US-0373817P.  
 19-APR-2002; 2002US-0373826P.  
 19-APR-2002; 2002US-0373884P.  
 22-APR-2002; 2002US-0374977P.  
 16-MAY-2002; 2002US-0381037P.  
 16-MAY-2002; 2002US-0381038P.  
 16-MAY-2002; 2002US-0381042P.  
 17-MAY-2002; 2002US-0381642P.  
 28-MAY-2002; 2002US-0383656P.  
 29-MAY-2002; 2002US-0383831P.  
 25-JUN-2002; 2002US-0391335P.  
 (SMIT/) SMITHSON G.  
 (MILL/) MILLET I.  
 (PEWM/) PEYMAN J A.  
 (KEKU/) KEKUDA R.  
 (JUOJ/) JU J.  
 (LILL/) LI L.  
 (GUOX/) GUO X.  
 (PATT/) PATTURAJAN M.  
 (SPYT/) SPYTEK K A.  
 (EDIN/) EDINGER S R.  
 (ELLE/) ELLERMAN K.  
 (MALY/) MALYANKAR U M.  
 (ORTT/) ORT T.  
 (GORM/) GORMAN L.  
 (ZERH/) ZERHUSEN B D.  
 (ANDE/) ANDERSON D W.  
 (ZHON/) ZHONG W.  
 (CATT/) CATTERTON E.  
 (JIWW/) JI W.  
 (MILL/) MILLER C E.  
 (RAST/) RASTELLI L.  
 (STON/) STONE D J.  
 (PENA/) PENA C E A.  
 (SHEN/) SHENOY S G.  
 (SHIM/) SHIMKETS R A.  
 (ROTH/) ROTHENBERG M E.  
 (LEAC/) LEACH M D.  
 (AGEE/) AGEE M L.  
 (BERG/) BERGHS C.  
 (DIP/) DIPIPPO V A.  
 (EISE/) EISEN A.  
 (GANG/) GANGOLLI E A.  
 (RIEG/) RIEGER D K.  
 (SPAD/) SPADERNA S K.  
 Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 WPI; 2004-213931/20.  
 N-PSDB; ADN62899.  
 Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 Claim 1; SEQ ID NO 96; 395pp; English.

xx The invention relates to isolated NOVX polypeptides and polynucleotides.  
 cc NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 cc treat a medical condition in human related to the aberrant expression and  
 cc activity of NOVX polypeptides. For example, NOVX polypeptides and  
 cc polynucleotides may be used to treat disorders associated with decreased  
 cc expression or activity of NOVX by supplementing the patient our  
 cc production or to rectify mutations. Conversely, antisense NA molecules  
 cc may be administered to down regulate expression of NOVX polypeptides by  
 cc binding with the cells own genes and preventing their expression. NOVX  
 cc polynucleotides and complementary sequences may also be used as DNA  
 cc probes in diagnostic assays to detect and quantitate the presence of  
 cc similar sequences in samples, and so which patients may be in need of  
 cc restorative therapy. NOVX polypeptides may also be used as antigens in  
 cc the production of antibodies and in assays to identify modulators  
 cc (agonists and antagonists) of the expression and activity of NOVX. The  
 cc anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 cc used to modulate NOVX polynucleotide expression and activity of NOVX  
 cc polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 cc diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 cc polypeptides and polynucleotides may be used in this way to prevent,  
 cc diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 cc disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 cc disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 cc haematopoietic disorders, and the various dyslipidaemias, metabolic  
 cc disturbances associated with obesity, the metabolic syndrome X and  
 cc wasting disorders associated with chronic diseases and various cancers.  
 cc They may also be used as antibacterial agents. The present sequence  
 cc represents the amino acid sequence of a human NOVX protein.  
 xx Sequence 198 AA;  
 SQ Query Match 100.0%; Score 41; DB 8; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 2.9; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;  
 Qy 1 MARSLLLPL 9  
 Db 1 MARSLLLPL 9  
 |||||  
 |||||  
 RESULT 6  
 AAR67888  
 ID AAR67888 standard; protein; 253 AA.  
 XX  
 AC AAR67888;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-AUG-1995 (first entry)  
 XX  
 XX Human stratum corneum chymotrophic recombinant enzyme (SCCE).  
 DE  
 XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;  
 KW callosities; keratosis pilaris; ichthyoses; eczema.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9500651-A1.  
 PN  
 XX  
 XX 05-JAN-1995.  
 PD  
 XX  
 XX 20-JUN-1994; 94WO-IB000166.  
 PF  
 XX  
 XX 18-JUN-1993; 93DK-00000725.  
 PR  
 XX (SYMB-) SYMBICOM AB.  
 PA  
 XX Egelrud T, Hansson L;  
 PI  
 XX  
 XX WPI; 1995-052088/07.  
 DR  
 XX N-PSDB; AAQ81203.  
 DR  
 XX Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and  
 PT

PT related vectors, transformed cells and polypeptides, useful for treating  
 PT skin disorders, e.g. acne or psoriasis, and for identification of  
 PT specific inhibitors.

PS Disclosure; Page 97; 137pp; English.

XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic  
 CC and skin care products, especially to treat and prevent acne, xeroderma,  
 CC or other hyperkeratotic conditions (e.g. callosities or keratosis  
 CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced  
 CC recombinantly following mammalian, insect, plant, or microorganism  
 CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN  
 CC field.)

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 DB 1 MARSLLLPL 9

RESULT 7  
 AAW05383  
 ID AAW05383 standard; protein; 253 AA.

XX AC AAW05383;

XX DT 31-DEC-1996 (first entry)

XX DE Human amyloid precursor protein protease.

XX KW Amyloid precursor protein protease; Alzheimer's disease; diagnosis;  
 KW therapy.

XX OS Homo sapiens.

XX PN WO9631122-A1.

XX PD 10-OCT-1996.

XX PF 02-APR-1996; 96WO-US004294.

XX PR 04-APR-1995; 95US-00416257.

XX PA (ELIL ) LILLY & CO ELI.

XX PI Dixon EP, Johnstone EM, Little SP;

XX DR WPI; 1996-464694/46.

XX DR N-PSDB; AAT39783.

XX PT New isolated human amyloid precursor protein protease - used to develop  
 PT prods. for the treatment or diagnosis of associated conditions, esp.  
 PT Alzheimer's disease.

XX PS Claim 1; Page 44-45; 55pp; English.

XX CC Human amyloid precursor protein protease (AAW05383) is involved in the  
 CC processing or clearance of amyloid precursor protein to form beta-amyloid  
 CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)  
 CC obtd. from a human lung library. Recombinant protease can be produced in  
 CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic  
 CC (partic. AV-120 host cells. It is used to develop products for the design  
 CC and testing of cpds. useful for treating or preventing conditions  
 CC associated with beta-amyloid peptide, esp. Alzheimer's disease

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 DB 1 MARSLLLPL 9

RESULT 8

ABB84421

ID ABB84421 standard; peptide; 253 AA.

XX AC ABB84421;

XX DT 08-NOV-2002 (first entry)

XX DE Human SCCE protein N-terminal fragment SEQ ID 48.

XX KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
 KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.

XX OS Homo sapiens.

XX PN WO2000262135-A2.

XX PD 15-AUG-2002.

XX PF 08-FEB-2002; 2002WO-IB001300.

XX PR 09-FEB-2001; 2001CA-02332855.

XX PR 09-FEB-2001; 2001DK-00000218.

XX PA (EGEL/) EGELRUD T.

XX PA (HANS/) HANSSON L.

XX PI Egelrud T, Hansson L;

XX DR WPI; 2002-643380/69.

XX PT Transgenic mammal or its embryo useful as model for human disease, has  
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
 PT enzyme operably linked to promoter that drives its expression in skin.

XX PS Example 6; Page 37; 74pp; English.

XX CC This invention describes a novel non-human transgenic mammal or mammalian  
 CC embryo having integrated within its genome, a heterologous nucleotide  
 CC sequence comprising at least a significant part of a nucleotide sequence  
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
 CC operably linked to a promoter that drives expression of heterologous scce  
 CC or its variant in skin. The product of the invention is useful as a model  
 CC for the study of disease with the aim of improving treatment, to relieve  
 CC or ameliorate a pathogenic condition, for development or testing of a  
 CC cosmetic or a pharmaceutical formulation, and for the development of a  
 CC diagnostic method. It can also be used as a model for a skin disease or  
 CC skin cancer. The invention is also useful for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of an  
 CC abnormal or unwanted phenotype, and for screening or identifying a  
 CC compound or composition effective for the prevention or treatment of  
 CC inflammatory skin diseases selected from diseases consisting of epidermal  
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful  
 CC as a model for further studies of itch mechanisms and the testing of  
 CC potential compounds and compositions for relieve of various skin diseases  
 CC where itch is a component. This sequence represents the N-terminal  
 CC fragment of the human stratum corneum chymotryptic enzyme, SCCE  
 CC synonymous with human kallikrein 7 (KLK7), used in the development of the  
 CC transgenic mammals described in the invention

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 5; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLLPL 9  
| | | | | | | | | |  
Db 1 MARSLLLLPL 9

RESULT 9  
ABB84406  
ID ABB84406 standard; protein; 253 AA.  
XX  
AC ABB84406;  
XX  
DT 08-NOV-2002 (first entry)  
XX  
DE Human SCCE protein.  
XX  
KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;  
KW serine protease; transgenic mammal; skin; skin disease; skin cancer;  
KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;  
KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.  
XX  
OS Homo sapiens.  
FN WO200262135-A2.  
XX  
PD 15-AUG-2002.  
XX  
PF 08-FEB-2002; 2002WO-IB001300.  
XX  
PR 09-FEB-2001; 2001CA-02332655.  
PR 09-FEB-2001; 2001DK-00000218.  
XX  
PA (EGEL/) EGELRUD T.  
PA (HANS/) HANSSON L.  
XX  
PI Egelrud T, Hansson L;  
XX  
WPI; 2002-643380/69.  
DR N-PSDB; ABQ76226.  
XX  
Transgenic mammal or its embryo useful as model for human disease, has  
PT heterologous nucleotide sequence coding for stratum corneum chymotryptic  
PT enzyme operably linked to promoter that drives its expression in skin.  
XX  
Claim 10; Page 58-59; 74pp; English.

This invention describes a novel non-human transgenic mammal or mammalian  
XX embryo having integrated within its genome, a heterologous nucleotide  
XX sequence comprising at least a significant part of a nucleotide sequence  
XX coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,  
XX operably linked to a promoter that drives expression of heterologous scce  
XX or its variant in skin. The product of the invention is useful as a model  
XX for the study of disease with the aim of improving treatment, to relieve  
XX or ameliorate a pathogenic condition, for development or testing of a  
XX cosmetic or a pharmaceutical formulation, and for the development of a  
XX diagnostic method. It can also be used as a model for a skin disease or  
XX skin cancer. The invention is also useful for screening or identifying a  
XX compound or composition effective for the prevention or treatment of an  
XX abnormal or unwanted phenotype, and for screening or identifying a  
XX compound or composition effective for the prevention or treatment of  
XX inflammatory skin diseases selected from diseases consisting of epidermal  
XX hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,  
XX pruritus, atopic dermatitis, eczema, acne and inherited skin diseases  
XX with epidermal hyperkeratosis. The mammal of the invention is also useful  
XX as a model for further studies of itch mechanisms and the testing of  
XX potential compounds and compositions for relieve of various skin diseases  
XX where itch is a component. This sequence represents the human stratum  
XX corneum chymotryptic enzyme, SCCE which is a serine protease synonymous  
XX with human kallikrein 7 (KLK7) and is used in the development of the

CC transgenic mammals described in the invention  
XX  
SQ Sequence 253 AA;  
Query Match 100.0%; Score 41; DB 5; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLLPL 9  
| | | | | | | | | |  
Db 1 MARSLLLLPL 9

RESULT 10  
AAU82740  
ID AAU82740 standard; protein; 253 AA.  
XX  
AC AAU82740;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE Amino acid sequence of novel human human protease #39.  
XX  
KW Human; protease; cancer; immune-related disorder; cardiovascular disease;  
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;  
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
KW ocular disease; cytostatic; enzyme.  
XX  
OS Homo sapiens.  
XX  
PN WO200200860-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 26-JUN-2001; 2001WO-US020171.  
XX  
PR 26-JUN-2000; 2000US-0214047P.  
XX  
PA (SUGE-) SUGEN INC.  
XX  
PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
PI Charyczak G;  
XX  
WPI; 2002-139913/18.  
DR N-PSDB; ABK31782.  
XX  
Nucleic acids encoding novel human proteases, useful for useful for  
PT treating diseases and disorders such as cancers, immune-related diseases  
PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory  
PT disorders.  
XX  
Claim 6; Fig 2N; 313pp; English.

The present invention relates to the isolation of novel human proteases,  
XX and the nucleic acids encoding them. The sequences of the invention are  
XX useful for treating diseases and disorders such as cancers (e.g. breast,  
XX colon, lung), immune-related diseases and disorders (e.g. inflammatory  
XX diseases and asthma), cardiovascular diseases (e.g. restenosis and  
XX coronary thrombosis), brain or neuronal-associated diseases, metabolic  
XX disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.  
XX rheumatoid arthritis and psoriasis), central or peripheral nervous system  
XX diseases, migraines, pain, sexual dysfunction, mood disorders, attention  
XX disorders, neurological disorders (e.g. Alzheimer's disease, Parkinson's  
XX disease) and dyskinesias. The nucleic acids and polypeptides are also  
XX useful for treating viral infections caused by human immunodeficiency  
XX virus (HIV), and non-viral infections such as ocular disease (e.g.  
XX glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel  
XX human proteases of the invention  
SQ Sequence 253 AA;



Query Match 100.0%; Score 41; DB 5; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3.8; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 DB 1 MARSLLLPL 9

RESULT 11  
 ABU07440  
 ID ABU07440 standard; protein; 253 AA.  
 AC ABU07440;  
 XX  
 DT 28-JAN-2003 (first entry)  
 DE Protein differentially regulated in prostate cancer #43.  
 KW Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX  
 OS Homo sapiens.  
 PN WO200281638-A2.  
 XX  
 PD 17-OCT-2002.  
 PF 08-APR-2002; 2002WO-US010824.  
 XX  
 PR 06-APR-2001; 2001US-0281731P.  
 PR 06-APR-2001; 2001US-0281732P.  
 XX  
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 PI Sun Z, Jay G;  
 XX WPI; 2003-058520/05.  
 DR N-PSDB; ABX10343.  
 XX

Novel genes which are differentially regulated in prostate cancer, useful  
 for diagnosing prostate cancer in prostate tissue sample and assessing  
 therapeutic or preventive intervention in prostate cancer patients.  
 Claim 1; Page 293-294; 416pp; English.

The invention describes genes (I) which are differentially regulated in  
 prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 sample comprising prostate tissue, which involves determining the number  
 of target genes which are differentially-regulated in the sample, where  
 the number is indicative of the probability that the sample comprises  
 prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 intervention in a subject having a prostate cancer, which involves  
 determining the expression levels in a sample comprising prostate tissue  
 of target genes which are differentially-regulated in prostate cancer.  
 Preferably, the expression levels of at least 10 genes are determined.  
 (I) is also useful for identifying agents that modulate a biological  
 activity of a polypeptide differentially-regulated in prostate cancer  
 cells, which involves contacting a polypeptide differentially-regulated  
 in prostate cancer cells with a test agent under conditions effective for  
 the test agent to modulate a biological activity of the polypeptide, and  
 determining whether the test agent modulates the biological activity. (I)  
 is useful as molecular markers, as drug targets, and for detecting.  
 (I) is useful for diagnosing, assessing, monitoring, prognosticating,  
 preventing or treating, determining predisposition to diseases and  
 conditions especially relating to prostate cancer. (I) and its expression  
 products are used in the diagnostic test to assay for presence of cancer  
 e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 of cancer, its stage of development, the nature of genetic defect, etc.  
 The polypeptide encoded by (I) can be used as target for therapy or drug

CC discovery. (I) can also be used for expressing the polypeptide and thus  
 for searching specific binding partners of the polypeptide. (I) is useful  
 in therapeutic applications to treat prostate cancer. The identification  
 of specific genes, and groups of genes, expressed in pathways  
 CC physiologically relevant to prostate cancer permits the definition of  
 CC functional and disease pathways and the delineation of targets in these  
 CC pathways which are useful in diagnostic, therapeutic, and clinical  
 CC applications. This is the amino acid sequence of a protein differentially  
 CC regulated in prostate cancer  
 XX

Query Match 100.0%; Score 41; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3.8; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 DB 1 MARSLLLPL 9

RESULT 12  
 ABU07471  
 ID ABU07471 standard; protein; 253 AA.  
 AC ABU07471;  
 XX  
 DT 28-JAN-2003 (first entry)  
 DE Protein differentially regulated in prostate cancer #74.  
 KW Prostate cancer; gene expression; differential regulation;  
 KW molecular marker; drug target; cancer detection; cancer diagnosis;  
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX  
 OS Homo sapiens.  
 PN WO200281638-A2.  
 XX  
 PD 17-OCT-2002.  
 PF 08-APR-2002; 2002WO-US010824.  
 XX  
 PR 06-APR-2001; 2001US-0281731P.  
 PR 06-APR-2001; 2001US-0281732P.  
 XX  
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 PI Sun Z, Jay G;  
 XX WPI; 2003-058520/05.  
 DR N-PSDB; ABX10375.  
 XX

Novel genes which are differentially regulated in prostate cancer, useful  
 for diagnosing prostate cancer in prostate tissue sample and assessing  
 therapeutic or preventive intervention in prostate cancer patients.  
 Claim 1; Page 351; 416pp; English.

The invention describes genes (I) which are differentially regulated in  
 prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
 sample comprising prostate tissue, which involves determining the number  
 of target genes which are differentially-regulated in the sample, where  
 the number is indicative of the probability that the sample comprises  
 prostate cancer. (I) is useful for assessing a therapeutic or preventive  
 intervention in a subject having a prostate cancer, which involves  
 determining the expression levels in a sample comprising prostate tissue  
 of target genes which are differentially-regulated in prostate cancer.  
 Preferably, the expression levels of at least 10 genes are determined.  
 (I) is also useful for identifying agents that modulate a biological  
 activity of a polypeptide differentially-regulated in prostate cancer  
 cells, which involves contacting a polypeptide differentially-regulated  
 in prostate cancer cells with a test agent under conditions effective for

CC the test agent to modulate a biological activity of the polypeptide, and  
 CC determining whether the test agent modulates the biological activity. (I)  
 CC is useful as molecular markers, as drug targets, and for detecting,  
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
 CC preventing or treating, determining predisposition to diseases and  
 CC conditions especially relating to prostate cancer. (I) and its expression  
 CC products are used in the diagnostic test to assay for presence of cancer  
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
 CC of cancer, its stage of development, the nature of genetic defect, etc.  
 CC The polypeptide encoded by (I) can be used as target for therapy or drug  
 CC discovery. (I) can also be used for expressing the polypeptide and thus  
 CC for searching specific binding partners of the polypeptide. (I) is useful  
 CC in therapeutic applications to treat prostate cancer. The identification  
 CC of specific genes, and groups of genes, expressed in pathways  
 CC physiologically relevant to prostate cancer permits the definition of  
 CC functional and disease pathways and the delineation of targets in these  
 CC pathways which are useful in diagnostic, therapeutic, and clinical  
 CC applications. This is the amino acid sequence of a protein differentially  
 CC regulated in prostate cancer

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MARSLLPL 9  
 |||||

Db 1 MARSLLPL 9

RESULT 13

ABR58471  
 ID ABR58471 standard; protein; 253 AA.

XX ABR58471;

XX 07-JUL-2003 (first entry)

XX Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.

XX Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.

XX Homo sapiens.

XX WO2003029468-A1.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031467.

XX 02-OCT-2001; 2001US-0327135P.

XX 30-MAY-2002; 2002US-0384531P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Mannion J;

XX WPI; 2003-372001/35.

XX New polynucleotide and polypeptide useful for diagnosing and/or treating  
 PT cancer, particularly ovarian cancer, and as a vaccine.

XX Claim 2; Page 157-158; 169pp; English.

XX The invention relates to a novel isolated polynucleotide. The  
 CC polynucleotides of the invention have cytostatic activity, and may have a  
 CC use in gene therapy, and in a vaccine. The composition and methods are  
 CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.  
 CC The composition may also be used as a vaccine to prevent cancer. The  
 CC present sequence is used in the exemplification of the invention

SQ Sequence 253 AA;

Query Match 100.0%; Score 41; DB 6; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MARSLLPL 9

Db 1 MARSLLPL 9

RESULT 14

ADB80484

ID ADB80484 standard; protein; 253 AA.

XX ADB80484;

XX 04-DEC-2003 (first entry)

XX Ovarian cancer-associated protein #24.

XX Cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;  
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;  
 KW pre-cancerous lesion detection.

XX Homo sapiens.

XX WO2002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

XX 27-AUG-2001; 2001US-0315287P.

XX 05-SEP-2001; 2001US-0317544P.

XX 13-NOV-2001; 2001US-0350666P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

XX N-PSDB; ADB80483.

XX Detecting an ovarian cancer-associated transcript in a cell from a  
 PT patient, comprises contacting a biological sample from the patient with a  
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX Claim 13; Page 291; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-  
 CC associated transcript in a cell from a patient, by contacting a  
 CC biological sample from the patient with a polynucleotide that selectively  
 CC hybridizes to a sequence at least 80% identical to any of one of 80  
 CC nucleic acid sequences given in the specification. The method is useful  
 CC in diagnosing ovarian cancer and in identifying and using agents and/or  
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,  
 CC polypeptide and the antibody may also be used in detecting ovarian  
 CC cancers, monitoring and early detection of relapse following treatment,  
 CC monitoring response to therapy, selecting patients for post-operative  
 CC chemotherapy or radiation therapy, in selecting mode of therapy,  
 CC determining tumour prognosis, early detection of pre-cancerous lesions,  
 CC and as vaccines. This sequence corresponds to one of the proteins used  
 CC for the detection method of the invention.

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 7; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
| | | | | | | |  
Db 1 MARSLLLPL 9

Search completed: July 13, 2005, 17:19:56  
Job time : 77.6667 secs

## RESULT 15

ADJ68833  
ID ADJ68833 standard; protein; 253 AA.

XX AC ADJ68833;

XX DT 06-MAY-2004 (first entry)

XX DE Human heat mitochondrial protein as a therapeutic target SeqID639.

XX KW mitochondrial; human; screening assay; diabetes mellitus;

XX KW Huntington's disease; osteoarthritis;

XX KW Leber's hereditary optic neuropathy; LHON;

XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

XX KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;

XX KW osteopathic; ophthalmological; cytostatic.

XX OS Homo sapiens.

XX PN WO2003087768-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010870.

XX PR 12-APR-2002; 2002US-0372843P.

XX PR 17-JUN-2002; 2002US-0389987P.

XX PR 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) MITOKOR.

XX PA (BUCK-) BUCK INST AGE RES.

XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

XX FI Warnock DE;

XX DR WPI; 2003-845369/78.

XX PT Identifying a mitochondrial target for drug screening assays and for  
PT treating diseases associated with altered mitochondrial function,  
PT comprises detecting a modified polypeptide in a sample and correlating  
PT with the disease.

PS Claim 1; SEQ ID NO 639; 180pp; English.

XX CC This invention relates to novel mitochondrial targets that can be used  
CC for therapeutic intervention in treating a disease associated with  
CC altered mitochondrial function. Specifically, it refers to a method for  
CC identifying proteins of the human heart mitochondrial proteome that are  
CC useful for drug screening assays, as well as therapeutic targets. The  
CC present invention describes a method for identifying such proteins that  
CC can be used in the treatment of various diseases associated with altered  
CC mitochondrial function including diabetes mellitus, Huntington's disease,  
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
CC compositions have neuroprotective, nontropic, antidiabetic,  
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
CC cytostatic activities. This polypeptide sequence is a human heart  
CC mitochondrial protein of the invention.

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 41; DB 7; Length 253;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9

This page Blank (usp10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083a-86

Perfect score: 41

Sequence: 1 MARSLLPL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	253	2 A53968	serine proteinase
2	34	82.9	461	2 A46394	suppressor protein
3	33	80.5	397	2 B87343	conserved hypothet
4	33	80.5	571	2 E96550	hypothetical prote
5	33	80.5	582	2 I48673	matrix metalloprot
6	32	80.5	582	2 I84471	matrix metalloprot
7	32	78.0	81	2 B97856	hypothetical prote
8	32	78.0	126	2 C82169	DNA-binding protei
9	32	78.0	432	2 A83060	hypothetical prote
10	32	78.0	506	2 B87102	conserved membrane
11	32	78.0	921	2 T51136	ionotropic glutama
12	32	78.0	923	2 F84732	probable ligand-ga
13	32	78.0	1628	2 T38055	hypothetical prote
14	31	75.6	138	2 I48107	sodium channel alp
15	31	75.6	147	2 S28698	hypothetical prote
16	31	75.6	169	2 A80627	cell division inhi
17	31	75.6	169	2 B29016	cell division inhi
18	31	75.6	294	2 B86450	hypothetical prote
19	31	75.6	346	2 T11364	NADH2 dehydrogenas
20	31	75.6	375	2 C71917	probable transamin
21	31	75.6	375	2 H64597	probable transamin
22	31	75.6	452	2 A83734	PTS system, n-acet
23	31	75.6	491	2 J06197	stromelysin 3 (EC
24	31	75.6	556	1 S31330	inulinase (EC 3.2.
25	31	75.6	708	2 T43109	cytolysin B transp
26	31	75.6	2005	2 B25019	sodium channel pro
27	31	75.6	5069	2 T17464	rifamycin polyketi
28	30	73.2	72	2 JC2384	corazonin precurs
29	30	73.2	74	2 T17834	hypothetical prote

30	30	73.2	170	2 S43476	histone-like DNA-b
31	30	73.2	216	2 J50297	DRE/CRT-binding pr
32	30	73.2	216	2 T51830	transcription fact
33	30	73.2	222	2 D82132	hypothetical prote
34	30	73.2	299	2 T17832	hypothetical prote
35	30	73.2	314	2 D85294	transcription acti
36	30	73.2	314	2 T05799	transcription acti
37	30	73.2	323	2 S47741	probable transcrip
38	30	73.2	323	2 B86025	probable transcrip
39	30	73.2	323	2 A91179	probable transcrip
40	30	73.2	325	2 A88452	protein ZC155.4 (I
41	30	73.2	325	2 T25122	hypothetical prote
42	30	73.2	325	2 A97482	hypothetical prote
43	30	73.2	325	2 A12699	conserved hypothet
44	30	73.2	331	2 G87494	hypothetical prote
45	30	73.2	354	2 A48931	transmembrane glyc

#### ALIGNMENTS

##### RESULT 1

A53968

serine proteinase SCCE precursor - human

N:Alternate names: stratum corneum chymotryptic enzyme

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A53968

R:Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A:Reference number: A53968; MUID:94308225; PMID:8034709

A:Accession: A53968

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <HAN>

A:Cross-references: UNIPROT:P49862; GB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G5325504

C:Genetics: GDB:PRSS6; SCCE

A:Cross-references: GDB:377730

A:Map position: 7q35-7q35

C:Superfamily: trypsin; trypsin homology

F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 41; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLPL 9  
| | | | | | | |  
DB 1 MARSLLPL 9

##### RESULT 2

A46394

suppressor protein SSU1 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein U1531; protein YLR005W

C:Species: Saccharomyces cerevisiae

C>Date: 18-May-1994 #sequence\_revision 19-Jul-1996 #text\_change 09-Jul-2004

C:Accession: A46394; S64827

R:Yoon, H.; Miller, S.P.; Pabich, E.K.; Donahue, T.F.

Genes Dev. 6, 2463-2477, 1992

A:Title: SSU1, a suppressor of a HIS4 5'-UTR stem-loop mutation, is essential for transla

A:Reference number: A46394; MUID:94040711; PMID:1340463

A:Accession: A46394

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-461 <YOO>

A:Cross-references: UNIPROT:Q04673; GB:Z17385; NID:G2695; PID:G2696

R:Vandenbol, M.; Portetelle, D.; Hilger, F.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64742

A:Accession: S64827

A:Molecule type: DNA  
A:Residues: 1-461 <VAN>  
A:Cross-references: EMBL:Z73177; NID:g1360293; PID:g1360294; MIPS:YLR005W  
A:Note: experimental\_source strain S288C  
C:Genetics8:  
A:Gene: SGD:SSL1  
A:Cross-references: SGD:80003995; MIPS:YLR005W  
A:Map position: 12R  
C:Superfamily: RNA polymerase II transcription initiation/nucleotide excision repair factor  
C:Keywords: transmembrane protein  
F:356-372/Domain: transmembrane #status predicted <TMM>

Query Match 82.9%; Score 34; DB 2; Length 461;  
Best Local Similarity 77.8%; Pred. No. 21;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 MARSLLLPL 9  
||| |||||  
Db 216 MARGLLPLV 224

RESULT 3  
B87343  
conserved hypothetical protein CC0757 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: B87343  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B. J. Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, N.; J. Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: B87343  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <STO>  
A:Cross-references: UNIPROT:Q9AA49; GB:AE005673; NID:g13421992; PIDN:AAK22742.1; GSPDB:G  
C:Genetics:  
A:Gene: CC0757

Query Match 80.5%; Score 33; DB 2; Length 397;  
Best Local Similarity 66.7%; Pred. No. 29;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
||| |||||  
Db 285 LARALLPL 293

RESULT 4  
E96550  
hypothetical protein F11M15.13 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: E96550  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huijzer, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: E96550  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-571 <STO>  
A:Cross-references: UNIPROT:Q9SYC9; GB:AE005173; NID:g4836937; PIDN:AAD30639.1; GSPDB:GN  
C:Genetics:

A:Gene: F11M15.13  
A:Map position: 1

Query Match 80.5%; Score 33; DB 2; Length 571;  
Best Local Similarity 77.8%; Pred. No. 42;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
||| |||||  
Db 332 MLRSLLVPL 340

RESULT 5  
I48673  
matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 22-Jun-1999  
C:Accession: I48673  
R:Okada, A.; Belloq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P. Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995  
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cells  
A:Reference number: I38046; MUID:95224014; PMID:7708715  
A:Accession: I48673  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-582 <RES>  
A:Cross-references: EMBL:X83536; NID:g804999; PIDN:CAA58520.1; PID:g805000  
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotein  
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen  
F:1-23/Domain: signal sequence #status predicted <Sig>  
F:24-97/Domain: activation peptide #status predicted <PRO>  
F:61-284/Domain: matrix metalloproteinase homology <MMP>  
F:313-508/Domain: hemopexin repeat homology <PXN>  
F:93-239, 243, 249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status I  
F:239, 243, 249/Binding site: zinc, catalytic (His) (active) #status predicted  
F:240/Active site: Glu #status predicted

Query Match 80.5%; Score 33; DB 2; Length 582;  
Best Local Similarity 87.5%; Pred. No. 43;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ARSLLLPL 9  
: |||||  
Db 8 SRSLLLPL 15

RESULT 6  
I84471  
matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - rat  
N:Alternate names: membrane-type metalloproteinase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I84471; I61946  
R:Okada, A.; Belloq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P. Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995  
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cells  
A:Reference number: I38046; MUID:95224014; PMID:7708715  
A:Accession: I84471  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-582 <RES>  
A:Cross-references: UNIPROT:Q10739; EMBL:X83537; NID:g805012; PIDN:CAA58521.1; PID:g805013  
A:Accession: I61946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-57, 'M', 69-254, 'A', 256-582 <RE2>  
A:Cross-references: EMBL:X91785; NID:g1001926; PIDN:CAA62897.1; PID:g1001927  
C:Genetics:  
A:Gene: mt-mmp  
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotein  
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen  
F:1-23/Domain: signal sequence #status predicted <Sig>  
F:24-97/Domain: activation peptide #status predicted <PRO>

F:61-284/Domain: matrix metalloproteinase homology <MMP>  
F:313-508/Domain: hemopexin repeat homology <PXN>  
F:93,239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status  
F:233,243,249/binding site: zinc, catalytic (His) (active) #status predicted  
F:240/Active site: Glu #status predicted

Query Match 80.5%; Score 33; DB 2; Length 582;  
Best Local Similarity 87.5%; Pred. No. 43;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ARSLLPL 9  
Db 8 SRSLPL 15

RESULT 7  
B97856  
hypothetical protein RC1250 [imported] - Rickettsia conorii (strain Malish 7)  
C:Species: Rickettsia conorii  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: B97856  
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R  
Science 293, 2093-2098, 2001  
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
A:Reference number: A97700; MUID:21442074; PMID:11557893  
A:Accession: B97856  
A:Status: preliminary  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-81 <KUR>  
A:Cross-references: UNIPROT:Q92G73; GB:AE006914; PIDN:AAL03788.1; PID:G15620385; GSPDB:C  
C:Genetics:  
A:Gene: RC1250

Query Match 78.0%; Score 32; DB 2; Length 81;  
Best Local Similarity 75.0%; Pred. No. 9.4;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ARSLLPL 9  
Db 43 ARSLLPI 50

RESULT 8  
C82169  
DNA-binding protein inhibitor Id-2-related protein VC1696 [imported] - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: C82169  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: C82169  
A:Status: preliminary  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-126 <HEI>  
A:Cross-references: UNIPROT:Q9KRE6; GB:AE004247; GB:AE003852; NID:G9656204; PIDN:AAF9484  
A:Experimental source: serogroup O1, strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1696  
A:Map position: 1

Query Match 78.0%; Score 32; DB 2; Length 126;  
Best Local Similarity 87.5%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MARSLLPL 8  
Db 33 MLRSLLP 40

## RESULT 9

AR3060  
hypothetical protein PA4684 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: AR3060  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bri  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: AR3060  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-432 <STO>  
A:Cross-references: UNIPROT:Q9HVB1; GB:AE004882; GB:AE004091; NID:G9950939; PIDN:AAG0807  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4684

Query Match 78.0%; Score 32; DB 2; Length 432;  
Best Local Similarity 87.5%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ARSLLPL 9  
Db 245 ARELLPL 252

## RESULT 10

B87102  
conserved membrane protein ML1544 [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: B87102  
R:Cole, S.T.; Bigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: B87102  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-506 <STO>  
A:Cross-references: UNIPROT:Q9Z5I3; GB:AL450380; NID:G13093364; PIDN:CAC30495.1; GSPDB:G  
C:Genetics:  
A:Gene: ML1544  
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1782

Query Match 78.0%; Score 32; DB 2; Length 506;  
Best Local Similarity 77.8%; Pred. No. 62;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MARSLLPL 9  
Db 223 MNRSLPL 231

## RESULT 11

TS1136  
ionotropic glutamate receptor glr5 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 09-Jul-2004  
C:Accession: TS1136  
R:Davenport, R.J.; Kiegle, E.A.; Tester, M.  
submitted to the EMBL Data Library, December 1999  
A:Description: GLR5, an ionotropic glutamate receptor ortholog from Arabidopsis.  
A:Reference number: Z25309  
A:Accession: TS1136  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-921 <DAV>  
A:Cross-references: UNIPROT:Q9SDQ4; EMBL:AF210701; PIDN:AAF21042.1  
A:Experimental source: Cultivar Columbus  
C:Genetics:  
A:Gene: glr5  
A:Map position: 2

Query Match 78.0%; Score 32; DB 2; Length 921;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
|||||:|  
Db 705 MARSRLVPL 713

## RESULT 12

F84732  
Probable ligand-gated ion channel subunit [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: F84732  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84732  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-923 <STO>  
A:Cross-references: UNIPROT:Q9SDQ4; GB:AE002093; NID:g3831456; PIDN:AAC69938.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g32400  
A:Map position: 2

Query Match 78.0%; Score 32; DB 2; Length 923;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
|||||:|  
Db 707 MARSRLVPL 715

## RESULT 13

T38055  
Hypothetical protein SPAC22F3.14C - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T38055; T38177; S64429  
R:Lyde, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, February 1995  
A:Reference number: Z21765  
A:Accession: T38055  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-26 <LY1>  
A:Cross-references: UNIPROT:Q09779; EMBL:Z69239; PIDN:CAA93223.1; GSPDB:GN00066; SPDB:SF  
R:Lyde, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: Z21776  
A:Accession: T38177  
A:Molecule type: DNA  
A:Residues: 8-1621 <LY2>  
A:Cross-references: EMBL:Z54285; NID:g1008429; PIDN:CAA91079.2; GSPDB:GN00066; SPDB:SPAC  
C:Genetics:  
A:Gene: SPAC1D4.14  
A:Map position: 1L

Query Match 78.0%; Score 32; DB 2; Length 1628;

Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RSLLLPL 9  
|||||  
Db 583 RSLLLPL 589

## RESULT 14

I48107  
sodium channel alpha subunit - long-tailed hamster (fragment)  
C:Species: Cricetus longicaudatus (long-tailed hamster)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I48107  
R:Lalik, P.H.; Krafte, D.S.; Ciccarelli, R.B.  
Am. J. Physiol. 264, 803-809, 1993  
A:Title: Characterization of endogenous sodium channel gene expressed in chinese hamster  
A:Reference number: I48107  
A:Accession: I48107  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-138 <RES>  
A:Cross-references: UNIPROT:Q60463; GB:M87540; NID:g191067; PIDN:AAA36978.1; PID:g191068  
C:Genetics:  
A:Gene: chol  
C:Superfamily: sodium channel protein  
C:Keywords: duplication

Query Match 75.6%; Score 31; DB 2; Length 138;  
Best Local Similarity 75.0%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLP 8  
|||||:|  
Db 1 MARSVLVP 8

## RESULT 15

S28698  
hypothetical protein 16 - Agrobacterium tumefaciens plasmid pTi15955  
C:Species: Agrobacterium tumefaciens  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C:Accession: S28698  
R:Barker, R.F.; Idler, K.B.; Thompson, D.V.; Kemp, J.D.  
Plant Mol. Biol. 2, 335-350, 1983  
A:Title: Nucleotide sequence of the T-DNA region from the Agrobacterium tumefaciens octo  
A:Reference number: S28683  
A:Accession: S28698  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-147 <BAR>  
A:Cross-references: UNIPROT:Q44395; EMBL:X00493; NID:g39062; PIDN:CAA25178.1; PID:g39078  
C:Genetics:  
A:Genome: plasmid

Query Match 75.6%; Score 31; DB 2; Length 147;  
Best Local Similarity 87.5%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLP 8  
|||||  
Db 1 MARVLLLP 8

Search completed: July 13, 2005, 17:31:32  
Job time : 14.5556 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-86  
Perfect score: 41  
Sequence: 1 MARSLLLPL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41	100.0	66	2	Q6DTY1	Q6dtl1 homo sapien
2	41	100.0	253	1	KLK7 HUMAN	P49862 homo sapien
3	41	100.0	253	2	Q8N5N9	Q8n5n9 homo sapien
4	35	85.4	396	2	Q9NZL1	Q9nzl1 homo sapien
5	35	85.4	468	2	Q9NZL2	Q9nzl2 homo sapien
6	35	85.4	518	1	BAE2 HUMAN	Q9y5z0 homo sapien
7	35	85.4	1135	2	Q84W49	Q84w49 arabidopsis
8	35	85.4	1192	2	Q9S7T0	Q9s7t0 arabidopsis
9	34	82.9	118	2	Q7RZ54	Q7rz54 neurospora
10	34	82.9	129	2	Q8H584	Q8h584 oryza sativ
11	34	82.9	322	2	Q67P07	Q67p07 symbiobacte
12	34	82.9	348	2	Q7UAK2	Q7uak2 synchococc
13	34	82.9	439	2	Q6PTA5	Q6pta5 candida gla
14	34	82.9	461	1	SSU1 YEAST	Q04673 saccharomyc
15	34	82.9	461	2	Q6B237	Q6b237 saccharomyc
16	33	80.5	188	2	Q6SSD7	Q6ssd7 triticum ae
17	33	80.5	272	2	Q8LC74	Q8lc74 arabidopsis
18	33	80.5	272	2	Q9FNC4	Q9fnc4 arabidopsis
19	33	80.5	295	2	Q8W2V5	Q8w2v5 oryza sativ
20	33	80.5	297	2	Q7P016	Q7p016 chromobacte
21	33	80.5	368	2	Q8SAT6	Q8sat6 oryza sativ
22	33	80.5	397	2	Q9AA49	Q9aa49 caulobacter
23	33	80.5	424	2	Q827F2	Q827f2 streptomyce
24	33	80.5	464	1	OTSA RHISN	P55612 rhizobium s
25	33	80.5	526	2	Q9ZRH9	Q9zrh9 oryza sativ
26	33	80.5	531	2	Q6ZK46	Q6zrk46 oryza sativ
27	33	80.5	571	2	Q9SYC9	Q9syc9 arabidopsis
28	33	80.5	582	1	MM14 MOUSE	P53690 mus musculus
29	33	80.5	582	1	MM14_RAT	Q10739 rattus norv
30	33	80.5	582	2	Q6DFU5	Q6dfu5 mus musculus
31	33	80.5	582	2	Q8BTX2	Q8btx2 mus musculus

32	33	80.5	809	2	Q8SS23	Q8se23 encephalito
33	33	80.5	934	2	Q9DER4	Q9der4 gallus gall
34	33	80.5	1048	2	Q6ZJF7	Q6zjf7 oryza sativ
35	33	80.5	1124	2	Q6ZBI6	Q6zbi6 oryza sativ
36	33	80.5	2998	2	Q8MXK6	Q8mxk6 leishmania
37	32	78.0	81	2	Q92G73	Q92g73 rickettsia
38	32	78.0	126	2	Q9KRE6	Q9kre6 vibrio chol
39	32	78.0	139	2	Q658A9	Q658a9 oryza sativ
40	32	78.0	149	2	Q9SYF9	Q9syf9 homo sapien
41	32	78.0	149	2	Q8KAK3	Q8kak3 chlorobium
42	32	78.0	178	2	Q6AHD8	Q6ahd8 leifsonia x
43	32	78.0	185	2	Q72AZ4	Q72az4 desulfovibr
44	32	78.0	226	2	Q8NH11	Q8nh11 homo sapien
45	32	78.0	227	2	Q9UKJ0	Q9ukj0 homo sapien

## ALIGNMENTS

RESULT 1  
Q6DTY1 PRELIMINARY; PRT; 66 AA.  
AC Q6DTY1;  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Kallikrein 7 splice variant 3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Kishi T., Michael I.P., Diamandis E.P.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AY646152; AAT56047.1; -  
SQ SEQUENCE 66 AA; 7171 MW; 82E1C392BC8222FDB CRC64;

Query Match 100.0%; Score 41; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 0.74; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Oy 1 MARSLLLPL 9  
Db 1 MARSLLLPL 9

RESULT 2  
KLK7 HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).  
GN Name=KLK7; Synonyms=PRSS6, SCCE;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
RC TISSUE=Skin;  
RX MEDLINE=94308225; PubMed=8034709;  
RA Hanson L., Stromqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;  
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";  
RL J. Biol. Chem. 269:19420-19426(1994).  
RN [2]  
RP SEQUENCE FROM N.A.

RC TISSUE=Keratinocytes;  
RA Yousef G.M., Scorilas A., Diamandis E.P.;  
RT "Molecular characterization, mapping and tissue expression of the  
RT human stratum corneum chymotryptic enzyme gene.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RP PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RX Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
RA Moss P., Paepel B., Wang K.;  
RA "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";  
RT Gene 257:119-130(2000).  
RN [4]  
RN SEQUENCE FROM N.A.  
RP Hanson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
RA Wallbrandt P., Egelrud T.;  
RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
RT mice; a model for chronic itchy dermatitis.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN CHARACTERIZATION.  
RX MEDLINE=95314630; PubMed=7794273;  
RA Skjott A., Stroemqvist M., Egelrud T.;  
RT "Primary substrate specificity of recombinant human stratum corneum  
RT chymotryptic enzyme.";  
RL Biochem. Biophys. Res. Commun. 211:586-589(1995).  
CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive  
CC structures in the cornified layer of the skin in the continuous  
CC shedding of cells from the skin surface. Specific for amino acid  
CC residues with aromatic side chains in the P1 position. SCCE  
CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
CC phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
CC activation of precursors to inflammatory cytokines.  
CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
CC also seen in the brain and kidney.  
CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L33404; AAC37551.1; -.  
DR EMBL; AF166330; AAD49718.1; -.  
DR EMBL; AF243527; AAG33360.1; -.  
DR EMBL; AF332583; AAK69624.1; -.  
DR PIR; A53968; A53968.  
DR HSSP; P00760; 1EZX.  
DR MEROPS; S01.300; -.  
DR Genew; HGNC:6368; KLK7.  
DR H-InVDB; HIX0015373; -.  
DR MIM; 604438; -.  
DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
DR GO; GO:0008544; P:epidermal differentiation; TAS.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYSP; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HTS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
KW Signal; Zymogen.  
FT SIGNAL 1 22  
FT PROPEP 23 29 Activation peptide.

FT CHAIN 30 253 Kallikrein 7.  
FT ACT\_SITE 70 70 Charge relay system (By similarity).  
FT ACT\_SITE 112 112 Charge relay system (By similarity).  
FT ACT\_SITE 205 205 Charge relay system (By similarity).  
FT DISULFID 36 137 By similarity.  
FT DISULFID 55 71 By similarity.  
FT DISULFID ? 239 By similarity.  
FT DISULFID 144 211 By similarity.  
FT DISULFID 176 190 By similarity.  
FT DISULFID 201 226 By similarity.  
FT CARBOHYD 246 246 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;  
Query Match 100.0%; Score 41; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MARSLLPL 9  
Db 1 MARSLLPL 9  
RESULT 3  
Q8N5N9 PRELIMINARY; PRT; 253 AA.  
AC Q8N5N9;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Stratum corneum chymotryptic enzyme, preproprotein.  
GN Name=KLK7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Skin;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins L.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hsieh F.,  
RA Datsenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,  
RA Brownstein M.J., Ustin T.B., Toshiki S., Carninci P., Carninci P.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; BC032005; AAH32005.1; -.  
DR HSSP; P00760; 1EZX.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.

DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00340; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 100.0%; Score 41; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 :|||||||  
 Db 1 MARSLLLPL 9

## RESULT 4

ID Q9NZL1 PRELIMINARY; PRT; 396 AA.  
 AC Q9NZL1;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Aspartyl protease.  
 GN Homo sapiens (Human).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20422477; PubMed=10965118;  
 RA Solans A., Estivill X., de La Luna S.;  
 RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to  
 RT Alzheimer's amyloid precursor protein beta-secretase.";  
 RL Cyogenet. Cell Genet. 89:177-184(2000).  
 CC -1- SIMILARITY: Belongs to peptidase family A1.  
 DR EMBL; AF188277; AAF35836.1; -;  
 DR HSSP; P56817; 1FKN.  
 DR GO; GO:0016021; C:integral to membrane; ISS.  
 DR GO; GO:0006509; P:membrane protein ectodomain proteolysis; ISS.  
 DR GO; GO:0042985; P:negative regulation of amyloid precursor pr. . .; ISS.  
 DR InterPro; IPR009119; Pept\_A1\_BACE.  
 DR InterPro; IPR009121; Pept\_A1\_BACE2.  
 DR InterPro; IPR009007; Pept\_Aspartic.  
 DR PRINTS; PR01817; BACE2.  
 DR PRINTS; PR01815; BACEFAMILY.  
 DR PROSITE; PS00141; ASP PROTEASE; 2.  
 KW Aspartyl protease; Hydrolase; Protease.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE 396 AA; 43013 MW; 5023A7AF391CBAC9 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 396;  
 Best Local Similarity 77.8%; Pred. No. 78;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 :|||||||  
 Db 4 LARALLPL 12

## RESULT 5

ID Q9NZL2 PRELIMINARY; PRT; 468 AA.  
 AC Q9NZL2;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 GN Name=BACE2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20422477; PubMed=10965118;  
 RA Solans A., Estivill X., de La Luna S.;  
 RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to  
 RT Alzheimer's amyloid precursor protein beta-secretase.";  
 RL Cyogenet. Cell Genet. 89:177-184(2000).  
 CC -1- SIMILARITY: Belongs to peptidase family A1.  
 DR EMBL; AF188276; AAF35835.1; -;  
 DR HSSP; P56817; 1FKN.  
 DR GO; GO:0016021; C:integral to membrane; ISS.  
 DR GO; GO:0006509; P:membrane protein ectodomain proteolysis; ISS.  
 DR GO; GO:0042985; P:negative regulation of amyloid precursor pr. . .; ISS.  
 DR GO; GO:0016486; P:peptide hormone processing; ISS.  
 DR InterPro; IPR001461; Peptidase A1.  
 DR InterPro; IPR009119; Pept\_A1\_BACE.  
 DR InterPro; IPR009121; Pept\_A1\_BACE2.  
 DR InterPro; IPR009007; Pept\_Aspartic.  
 DR InterPro; IPR001969; Pept\_Asp\_AS.  
 DR PRINTS; PR01817; BACE2.  
 DR PRINTS; PR01815; BACEFAMILY.  
 DR PROSITE; PS00141; ASP PROTEASE; 2.  
 KW Aspartyl protease; Hydrolase; Protease.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 468;  
 Best Local Similarity 77.8%; Pred. No. 92;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 :|||||||  
 Db 4 LARALLPL 12

## RESULT 6

BAE2\_HUMAN  
 ID BAE2\_HUMAN STANDARD; PRT; 518 AA.  
 AC Q9V5Z0; Q9UJ76;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Beta secretase 2 precursor (EC 3.4.23.45) (Beta-site APP-cleaving  
 DE enzyme 2) (Aspartyl protease 1) (Asp 1) (ASPI) (Membrane-associated  
 DE aspartic protease 1) (Memapsin-1) (Down region aspartic protease)  
 DE (UNQ418/PRO852).  
 GN Name=BACE2; Synonyms=ASP21;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057170; PubMed=10591213; DOI=10.1038/990107;  
 RA Yan R., Bienkowski M.J., Snuck M.E., Miso H., Tory M.C., Pauley A.M.,  
 RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,  
 RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.E.;  
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease beta-  
 RT secretase activity.";  
 RL Nature 402:533-537(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RA Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,  
 RA Giese K.;  
 RT "Identification of a novel aspartic-like protease differentially

expressed in human breast cancer cell lines.";   
 [3] Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.   
 RP SEQUENCE FROM N.A.   
 RA Accarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;   
 RT "Cloning of a gene from chromosome 21 Down region encoding a potential   
 RT transmembrane aspartyl protease.";   
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.   
 [4]   
 RP SEQUENCE FROM N.A.   
 RA Solans A., Estivill X., de la Luna S.;   
 RT "Cloning of a novel mammalian aspartyl protease.";   
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.   
 [5]   
 RP SEQUENCE FROM N.A.   
 RX MEDLINE=20120043; PubMed=10656250; DOI=10.1006/mcne.1999.0811;   
 RA Husain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,   
 RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,   
 RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;   
 RT "Identification of a novel aspartic proteinase (Asp 2) as beta-   
 RT secretase.";   
 RL Mol. Cell. Neurosci. 14:419-427 (1999).   
 [6]   
 RP SEQUENCE FROM N.A.   
 RX MEDLINE=20144060; PubMed=10677483; DOI=10.1073/pnas.97.4.1456;   
 RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;   
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of   
 RT beta-amyloid precursor protein.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460 (2000).   
 [7]   
 RP SEQUENCE FROM N.A.   
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;   
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,   
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,   
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,   
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,   
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,   
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,   
 RA Vandlen R., Watanabe C., Weiland D., Woods K., Xie M.-H., Yansura D.,   
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,   
 RA Godowski P., Gray A.;   
 RT "The secreted protein discovery initiative (SPDI), a large-scale   
 RT effort to identify novel human secreted and transmembrane proteins: a   
 RT bioinformatics assessment.";   
 RL Genome Res. 13:2265-2270 (2003).   
 [8]   
 RP SEQUENCE FROM N.A.   
 RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;   
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,   
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,   
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,   
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,   
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,   
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,   
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,   
 RA Minoshima S., Shimizu N., Nordieck G., Hornischer K., Brandt P.,   
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,   
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,   
 RA Wehmeyer S., Borzym K., Gardner K., Nizetic D., Francis F.,   
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;   
 RT "The DNA sequence of human chromosome 21.";   
 RL Nature 405:311-319 (2000).   
 [9]   
 RP SEQUENCE FROM N.A.   
 RC TISSUE=Skin;   
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;   
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,   
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,   
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,   
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,   
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,   
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,   
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";   
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).   
RN [10]   
RP CHARACTERIZATION.   
RX MEDLINE=22088158; PubMed=12093293; DOI=10.1021/bi025926t;   
RA Turner R.T. III, Loy J.A., Nguyen C., Devasamudram T., Ghosh A.K.,  
RA Koelsch G., Tang J.;  
RT "Specificity of memapsin 1 and its implications on the design of  
RT memapsin 2 (beta-secretase) inhibitor selectivity.";   
RL Biochemistry 41:8742-8746 (2002).   
CC -|- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-  
CC Val-Asn-Leu|-Asp-Ala-Glu-Phe in the Swedish variant of  
CC Alzheimer's amyloid precursor protein.   
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.   
CC -|- SIMILARITY: Belongs to the peptidase A1 family.   
CC -----   
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).   
CC -----   
DR EMBL; AF200342; AAF17078.1; -   
DR EMBL; AF117892; AAD45240.1; -   
DR EMBL; AF050171; AAD45963.1; -   
DR EMBL; AF178532; AAF29494.1; -   
DR EMBL; AF204944; AAF26368.1; -   
DR EMBL; AF200192; AAF13714.1; -   
DR EMBL; AY358927; AAQ89286.1; -   
DR EMBL; AL163284; CAB90458.1; -   
DR EMBL; AL163285; CAB90554.1; -   
DR EMBL; BC014453; AAH14453.1; -   
DR HSSP; P56817; 1MAH.   
DR MEROPS; A01.041; -   
DR Genew; HGNC:934; BACE2.   
DR MIM; 605668; -   
DR GO; GO:0005624; C:membrane fraction; TAS.   
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; TAS.   
DR GO; GO:0006464; P:protein modification; TAS.   
DR GO; GO:0009306; P:protein secretion; TAS.   
DR InterPro; IPR009119; Pept\_A1\_BACE.   
DR InterPro; IPR009121; Pept\_A1\_BACE2.   
DR InterPro; IPR001969; Pept\_Asp\_AS.   
DR InterPro; IPR009007; Pept\_Aspartic.   
DR InterPro; IPR001461; Peptidase\_A1.   
DR Pfam; PF00026; Asp; 1.   
DR PRINTS; PR01817; BACE2.   
DR PRINTS; PR01815; BACEFAMILY.   
DR PRINTS; PR00792; PEPSIN.   
DR PROSITE; PS00141; ASP\_PROTEASE; 2.   
KW Aspartyl protease; Glycoprotein; Hydrolase; Signal; Transmembrane;  
Zn-met.   
FT SIGNAL 1 20 Potential.   
FT PROPEP 21 ? Potential.   
FT CHAIN ? 518 Beta secretase 2.   
FT DOMAIN 21 473 Extracellular (Potential).   
FT TRANSMEM 474 494 Potential.   
FT DOMAIN 495 518 Cytoplasmic (Potential).   
FT ACT\_SITE 110 110 By similarity.   
FT ACT\_SITE 303 303 By similarity.

FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 366 366 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 36 36 A -> T (in Ref. 6).  
 SQ SEQUENCE 518 AA; 56180 MW; 2E903150823760D3 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 518;  
 Best Local Similarity 77.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 Db 4 LARALLPL 12

RESULT 7  
 Q84W49 PRELIMINARY; PRT; 1135 AA.  
 AC Q84W49;  
 DT 01-JUN-2003 (TREMELrel. 24, Created)  
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
 DE Hypothetical protein At3g01780 (Fragment).  
 GN Name=At3g01780;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsid.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Heuan V.W., Lee J.M.,  
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,  
 RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kaniya A., Kawai J., Kim C.J., Narusaka M.,  
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
 RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BT004227; AAC042242.1; -.  
 DR InterPro; IPR008938; ARM.  
 DR PROSITE; PS00018; EF HAND; UNKNOWN 1.  
 KW Hypothetical protein.  
 FT NON\_TER 1135 1135  
 SQ SEQUENCE 1135 AA; 126119 MW; C5FDDC178D1E2D96 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 1135;  
 Best Local Similarity 77.8%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 Db 237 MARSLVLPV 245

RESULT 8  
 Q9S7T0 PRELIMINARY; PRT; 1192 AA.  
 AC Q9S7T0;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)  
 DE F28J7.11 protein (F4P13.33 protein).  
 GN Name=F28J7.11; Synonyms=F4P13.33;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsid.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,  
 RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,

QY 2 ARSLLLPL 9  
 Db 237 MARSLVLPV 245

RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC010797; AAF03433.1; -.  
 DR EMBL; AC009325; AAF01560.1; -.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR002048; EF-hand.  
 DR PROSITE; PS00018; EF HAND; UNKNOWN 1.  
 SQ SEQUENCE 1192 AA; 132863 MW; 4F67B124CBAPF154 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 1192;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
 Db 237 MARSLVLPV 245

RESULT 9  
 Q7RZ54 PRELIMINARY; PRT; 118 AA.  
 AC Q7RZ54;  
 DT 01-MAR-2004 (TREMELrel. 26, Created)  
 DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)  
 DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)  
 DE Predicted protein (Hypothetical protein G21B4.400).  
 GN Name=NCU04420.1; Synonyms=G21B4.400;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Seditrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kamysellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,  
 RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,  
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nuebaum C., Birren B.;  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0:0-0(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AABX01000720; EAA28261.1; -.  
 DR EMBL; BX908808; CAF06025.1; -.  
 DR Hypothetical protein.  
 KW Hypothetical protein.  
 SQ SEQUENCE 118 AA; 12596 MW; DIF84E47108B2145 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 118;  
 Best Local Similarity 87.5%; Pred. No. 40;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARSLLLPL 9  
 Db 237 MARSLVLPV 245

Db 23 ARSLLPLM 30

## RESULT 10

Q8H584 PRELIMINARY; PRT; 129 AA.  
AC Q8H584;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein OJ1656\_E11.113 (Hypothetical protein P0496D04.53).  
GN Name=OJ1656\_E11.113; Synonyms=P0496D04.53;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Sasaki T., Matsumoto T., Yamamoto K.;  
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC clone: OJ1656\_E11.1";  
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Sasaki T., Matsumoto T., Yamamoto K.;  
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC clone: P0496D04.53";  
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003843; BAC24850.1; -;  
DR EMBL; AP004670; BAD30793.1; -;  
DR Gramene; O8H584; -;  
KW Hypothetical protein.  
SQ SEQUENCE 129 AA; 13891 MW; 6334E639E1A6DEAE CRC64;

Query Match 82.9%; Score 34; DB 2; Length 129;  
Best Local Similarity 87.5%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLPL 8  
|:|||||  
DB 40 LARSLLPL 47

## RESULT 11

Q67P07 PRELIMINARY; PRT; 322 AA.  
AC Q67P07;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Ferrichrome ABC transporter permease protein.  
GN ORFNames=STH1601;  
OS Symbiobacterium thermophilum.  
OC Bacteria; Actinobacteria; Symbiobacterium.  
OX NCBI\_TaxID=2734;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IAM14863;  
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T., Morimura K., Ikeda H., Hattori M., Beppu T.;  
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium thermophilum".  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP006840; BAD40586.1; -;  
DR InterPro; IPR000522; FecD.  
DR Pfam; PF01032; FecCD; 1.  
SQ SEQUENCE 322 AA; 34949 MW; 80157DE1532812D9 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 322;  
Best Local Similarity 77.8%; Pred. No. 16+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLPL 9  
|:|||||  
DB 1 MKRSLLPL 9

## RESULT 12

Q7U4K2 PRELIMINARY; PRT; 348 AA.  
AC Q7U4K2;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Putative ldpA protein.  
GN OrderedLocusNames=SYNW2065;  
OS Synecococcus sp. (strain WH8102).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.  
OX NCBI\_TaxID=84588;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;  
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L., Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarran J., Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;  
RT "The genome of a motile marine Synecococcus";  
RL Nature 424:1037-1042(2003).  
DR EMBL; BX569694; CAE08580.1; -;  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0005506; F:iron ion binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR001450; 4Fe4S\_ferredoxin.  
DR Pfam; PF00037; Fer4; 1.  
DR PROSITE; PS00198; 4FE4S\_FERREDOXIN; 1.  
KW 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding.  
SQ SEQUENCE 348 AA; 37531 MW; 1E1903E684081A95 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 348;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLPL 9  
|:|||||  
DB 298 MARSLLPL 306

## RESULT 13

Q6FTA5 PRELIMINARY; PRT; 439 AA.  
AC Q6FTA5;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Similar to sp|Q04673|Saccharomyces cerevisiae YLR005w SSL1.  
GN ORFNames=CAGL0G04059g;  
OS Candida glabrata CBS138.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=284593;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBS138;  
RG Genolevures;  
RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S., Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E., Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V., Barnay S., Blanchin L., Beckerich J.N., Beyne E., Bleykasten C., Boissrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A., Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A., Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O., Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A., Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,

```

RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissbach J.,
RA Winker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380953; CAG59466.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0004549; P:regulation of transcription; IEA.
DR InterPro; IPR004595; Ssl1.
DR InterPro; IPR007198; Ssl1 like.
DR InterPro; IPR002035; Vwf A.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF04056; Ssl1; 1.
DR SMART; SM00327; VWA; 1.
DR TIGRFAMs; TIGR00622; ssl1; 1.
DR PROSITE; PS02334; VMPA; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN 1.
SQ SEQUENCE 439 AA, 49183 MW, 07D1A63D1575D894 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 439;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 194 MARGLLLPV 202

RESULT 14
SSLL YEAST
ID SSLL YEAST STANDARD; PRT; 461 AA.
AC Q04673;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Suppressor of stem-loop protein 1.
GN Name=SSL1; OrderedLocustNames=YLR005W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94040711; PubMed=1340463;
RA Yoon H., Miller S.P., Pabich E.K., Donahue T.F.;
RT "SSL1, a suppressor of a His4 5'-UTR stem-loop mutation, is essential
RT for translation initiation and affects UV resistance in yeast.";
RL Genes Dev. 6:2463-2477(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoerge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Netwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetalle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestazu L.A., Vandenbol M., Verhasselt P.,
RA Vierdeldeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
RN [3]
RP SEQUENCE OF 400-461 FROM N.A.
RX MEDLINE=94239498; PubMed=8183345; DOI=10.1038/369242a0;
RA Maeda T., Wurgler-Murphy S.M., Saito H.;
RT "A two-component system that regulates an osmosensing MAP kinase
RT cascade in yeast.";

```

---

```

RL Nature 369:242-245(1994).
CC -!- FUNCTION: Essential for translation initiation and affects UV-
CC resistance in yeast.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z17385; CAA78992.1; -.
DR EMBL; Z73177; CAA97527.1; -.
DR EMBL; L26523; AAA35101.1; -.
DR PIR; A46394; A46394.
DR HSSP; Q13888; 1E53.
DR GerMOnline; 142067; -.
DR TRANSFAC; T02191; -.
DR SGD; S000003995; SSL1.
DR InterPro; IPR004595; Ssl1.
DR InterPro; IPR007198; Ssl1 like.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF04056; Ssl1; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR TIGRFAMs; TIGR00622; ssl1; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; FALSE_NEG.
KW Metal-binding; Nuclear protein; Zinc-finger.
FT ZN FING 427 449 C2H2-type.
SQ SEQUENCE 461 AA, 52290 MW, 57ADC630B790B4F CRC64;

Query Match 82.9%; Score 34; DB 1; Length 461;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 216 MARGLLLPV 224

RESULT 15
Q6B237
ID Q6B237 PRELIMINARY; PRT; 461 AA.
AC Q6B237;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE YLR005W.
GN Name=SSL1;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Margischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E.,
RA Hu Y., Vamberg F., Weger J., Kramer J., Moreira D., Kelley F.,
RA Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A.,
RA Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E.,
RA LaBaer J.;
RT "Creation of the YLEX clone resource: cloning of Saccharomycetes
RT cerevisiae ORFs in the Gateway recombinational cloning system.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY62893; AAT92912.1; -.
DR SGD; S000003995; SSL1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0008270; F;zinc ion binding; IEA.

```

DR GO; GO:0006281; P:DNA repair; IEA.  
DR GO; GO:0045449; P:regulation of transcription; IEA.  
DR InterPro; IPR004595; Ssl1.  
DR InterPro; IPR007198; Ssl1.  
DR InterPro; IPR002035; VWF\_A.  
DR InterPro; IPR007087; VWF\_C2H2.  
DR Pfam; PF04056; Ssl1; 1.  
DR SMART; SM00327; VWA; 1.  
DR TIGRFAMs; TIGR00622; Ssl1; 1.  
DR PROSITE; PS50234; VWEA; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
SQ SEQUENCE 461 AA; 52232 MW; 56DF9762644AF520 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 461;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
Db ||| |||:  
216 MARGLLLPV 224

Search completed: July 13, 2005, 17:29:29  
Job time : 64.2222 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-86

Perfect score: 41

Sequence: 1 MARSLLPL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

- Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pap:\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pap:\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pap:\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pap:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pap:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	3	US-09-502-600-86
2	41	100.0	9	4	US-09-918-243-86
3	41	100.0	253	2	US-08-557-146-2
4	41	100.0	253	2	US-08-824-874-3
5	41	100.0	253	2	US-09-154-344-2
6	41	100.0	253	3	US-08-930-188-2
7	41	100.0	253	3	US-09-210-084-3
8	41	100.0	253	4	US-09-764-762-3
9	41	100.0	253	5	PCT-US96-04294-2
10	41	100.0	265	4	US-09-943-016-7716
11	36	87.8	9	3	US-09-502-600-116
12	36	87.8	9	4	US-09-918-243-116
13	35	85.4	518	3	US-08-999-723-2
14	35	85.4	518	3	US-09-434-427-2
15	35	85.4	518	4	US-09-548-372D-2
16	35	85.4	518	4	US-09-548-367D-2
17	35	85.4	518	4	US-09-551-853D-2
18	35	85.4	518	4	US-09-215-450-19
19	35	85.4	518	4	US-09-416-901B-2
20	35	85.4	518	4	US-09-548-376D-2
21	35	85.4	518	4	US-09-886-143-2
22	35	85.4	518	4	US-09-794-927A-2
23	35	85.4	518	4	US-09-548-373D-2
24	35	85.4	518	4	US-09-795-847B-2
25	35	85.4	518	4	US-09-869-414-2
26	35	85.4	518	4	US-09-548-366F-2
27	35	85.4	518	4	US-09-548-368D-2

28	35	85.4	518	4	US-09-794-925A-2	Sequence 2, Appli
29	35	85.4	518	4	US-09-806-194A-2	Sequence 2, Appli
30	35	85.4	541	4	US-09-949-016-11075	Sequence 11075, A
31	32	78.0	43	4	US-09-149-476-355	Sequence 355, App
32	32	78.0	123	4	US-09-270-767-33977	Sequence 33977, A
33	32	78.0	123	4	US-09-270-767-49194	Sequence 49194, A
34	32	78.0	175	4	US-09-869-388-8	Sequence 8, Appli
35	32	78.0	204	4	US-09-134-000C-3659	Sequence 3659, Ap
36	32	78.0	226	4	US-09-869-388-10	Sequence 10, Appli
37	32	78.0	227	4	US-09-869-388-6	Sequence 6, Appli
38	32	78.0	230	4	US-09-869-388-4	Sequence 4, Appli
39	32	78.0	238	4	US-09-149-476-485	Sequence 485, App
40	32	78.0	303	3	US-08-985-950-2	Sequence 2, Appli
41	32	78.0	303	4	US-09-546-049-2	Sequence 2, Appli
42	32	78.0	303	4	US-09-869-388-2	Sequence 2, Appli
43	32	78.0	493	4	US-09-252-991A-30722	Sequence 30722, A
44	32	78.0	501	4	US-09-902-540-12135	Sequence 12135, A
45	32	78.0	654	4	US-09-907-794A-177	Sequence 177, App

## ALIGNMENTS

### RESULT 1

US-09-502-600-86  
; Sequence 86, Application US/09502600A  
; Patent No. 6294344  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; TITLE OF INVENTION: Ovarian Cancer  
; FILE REFERENCE: D6223CIP-C  
; CURRENT FILING DATE: 2000-02-11  
; CURRENT APPLICATION NUMBER: US/09/502,600A  
; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 86  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 1-9 of the SCCE protein  
US-09-502-600-86

Query Match 100.0%; Score 41; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLPL 9  
DB 1 MARSLLPL 9

### RESULT 2

US-09-918-243-86  
; Sequence 86, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 86  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens

```
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 1-9 of the SCCE protein
US-09-918-243-86

Query Match      100.0%; Score 41; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 3
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2

Query Match      100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 4
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-08-824-874-3

Query Match      100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 5
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-154-344-2

Query Match 100.0%; Score 41; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.45; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 MARSLLLPL 9  
Db 1 MARSLLLPL 9

RESULT 6  
US-08-930-188-2  
Sequence 2, Application US/08930188  
Patent No. 6093397  
GENERAL INFORMATION:  
APPLICANT: Dixon, Eric P.  
ADDRESSEE: Johnstone, Edward M.  
APPLICANT: Little, Sheila P.  
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND  
RELATED NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930,188  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/416,257  
FILING DATE: 04-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Blalock, Donna K.  
REGISTRATION NUMBER: 38,082  
REFERENCE/DOCKET NUMBER: X9239  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-277-1090  
TELEFAX: 317-276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-930-188-2

Query Match 100.0%; Score 41; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.45; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 MARSLLLPL 9  
Db 1 MARSLLLPL 9

RESULT 7  
US-09-210-084-3  
Sequence 3, Application US/09210084  
Patent No. 6197511  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,084  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,874  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0252 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 532504  
US-09-210-084-3

Query Match 100.0%; Score 41; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.45; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 MARSLLLPL 9  
Db 1 MARSLLLPL 9

RESULT 8  
US-09-764-762-3  
Sequence 3, Application US/09764762  
Patent No. 6472195  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Lal, Preeti  
TITLE OF INVENTION: NOVEL KALLIKREIN  
NUMBER OF SEQUENCES: 5

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

      Query Match      100.0%; Score 41; DB 4; Length 253;
      Best Local Similarity 100.0%; Pred. No. 0.45;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MARSLLLPL 9
Db      1 MARSLLLPL 9

RESULT 9
US-09-949-016-7716
; Sequence 7716, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7716
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Human
;
US-09-949-016-7716

      Query Match      100.0%; Score 41; DB 4; Length 265;
      Best Local Similarity 100.0%; Pred. No. 0.47;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MARSLLLPL 9
Db      13 MARSLLLPL 21

RESULT 11
US-09-502-600-116
; Sequence 116, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
```

; PRIOR APPLICATION NUMBER: 09/039,211  
; PRIOR FILING DATE: 03-14-1998  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 116  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 2-10 of the SCCE protein  
US-09-502-600-116

Query Match 87.8%; Score 36; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARSLLLPL 9  
Db 1 ARSLLLPL 8

RESULT 12  
US-09-918-243-116  
; Sequence 116, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 116  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 2-10 of the SCCE protein  
US-09-918-243-116

Query Match 87.8%; Score 36; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARSLLLPL 9  
Db 1 ARSLLLPL 8

RESULT 13  
US-08-999-723-2  
; Sequence 2, Application US/08999723A  
; Patent No. 6025180  
; GENERAL INFORMATION:  
; APPLICANT: Powell, David J.  
; APPLICANT: Southan, Christopher  
; APPLICANT: Chapman, Conrad G.  
; APPLICANT: Evans, Joanne R.  
; TITLE OF INVENTION: ASP1  
; FILE REFERENCE: GH70262  
; CURRENT APPLICATION NUMBER: US/08/999,723A  
; CURRENT FILING DATE: 1997-10-06  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-999-723-2

Query Match 85.4%; Score 35; DB 3; Length 518;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
Db 4 LARALLLPL 12

RESULT 14  
US-09-434-427-2  
; Sequence 2, Application US/09434427  
; Patent No. 6162630  
; GENERAL INFORMATION:  
; APPLICANT: POWELL, DAVID J.  
; APPLICANT: SOUTHAN, CHRISTOPHER  
; APPLICANT: CHAPMAN, CONRAD G.  
; APPLICANT: EVANS, JOANNE R.  
; TITLE OF INVENTION: ASP1  
; FILE REFERENCE: GH-70262-D1  
; CURRENT APPLICATION NUMBER: US/09/434,427  
; CURRENT FILING DATE: 1999-11-04  
; EARLIER APPLICATION NUMBER: US 08/999,723  
; EARLIER FILING DATE: 1997-10-06  
; EARLIER APPLICATION NUMBER: UK 9626022.9  
; EARLIER FILING DATE: 1996-12-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-434-427-2

Query Match 85.4%; Score 35; DB 3; Length 518;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9  
Db 4 LARALLLPL 12

RESULT 15  
US-09-548-372D-2  
; Sequence 2, Application US/09548372D  
; Patent No. 6420534  
; GENERAL INFORMATION:  
; APPLICANT: GURNEY ET AL.  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 29915/62801  
; CURRENT APPLICATION NUMBER: US/09/548,372D  
; CURRENT FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: US 60/155,493  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US 09/404,133  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: PCT/US99/20881  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US 60/101,594  
; PRIOR FILING DATE: 1998-09-24  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-548-372D-2

Query Match 85.4%; Score 35; DB 4; Length 518;  
Best Local Similarity 77.8%; Pred. No. 20;

	Matches	7;	Conservative	2;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MARSLLPL	9							
		:  :								
Db	4	LARALLPL	12							

Search completed: July 13, 2005, 17:34:29  
Job time : 20.4444 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds  
(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-86

Perfect score: 41

Sequence: 1 MARSLLPL 9

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1756220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	9 US-09-918-243-86	Sequence 86, Appl
2	41	100.0	9	9 US-09-905-083-86	Sequence 86, Appl
3	41	100.0	9	15 US-10-372-521-86	Sequence 85, Appl
4	41	100.0	9	15 US-10-831-075-86	Sequence 85, Appl
5	41	100.0	198	15 US-10-282-511-96	Sequence 96, Appl
6	41	100.0	253	9 US-09-888-615-98	Sequence 98, Appl
7	41	100.0	253	9 US-09-764-762-3	Sequence 3, Appl
8	41	100.0	253	14 US-10-071-214-2	Sequence 2, Appl
9	41	100.0	253	14 US-10-071-214-48	Sequence 48, Appl
10	41	100.0	253	14 US-10-264-283-90	Sequence 98, Appl
11	41	100.0	253	15 US-10-295-027-498	Sequence 40, Appl

12	41	100.0	253	15	US-10-173-999-48	Sequence 48, Appl
13	41	100.0	253	16	US-10-408-765A-639	Sequence 639, Appl
14	41	100.0	253	16	US-10-643-795A-95	Sequence 95, Appl
15	41	100.0	253	17	US-10-948-518-95	Sequence 95, Appl
16	41	100.0	253	17	US-10-868-490A-1	Sequence 1, Appl
17	41	100.0	257	15	US-10-344-394-38	Sequence 38, Appl
18	37	90.2	752	17	US-10-612-466B-6	Sequence 6, Appl
19	36	87.8	9	9	US-09-918-243-116	Sequence 116, Appl
20	36	87.8	9	9	US-09-905-083-116	Sequence 116, Appl
21	36	87.8	9	15	US-10-372-521-116	Sequence 116, Appl
22	36	87.8	9	16	US-10-831-075-116	Sequence 116, Appl
23	36	87.8	99	15	US-10-424-599-279409	Sequence 279409, A
24	36	87.8	172	15	US-10-424-599-239079	Sequence 239079, A
25	36	87.8	210	15	US-10-425-114-50699	Sequence 50699, A
26	36	87.8	542	16	US-10-425-115-318442	Sequence 318442, A
27	36	87.8	567	15	US-10-425-114-66721	Sequence 66721, A
28	36	87.8	818	9	US-09-888-615-111	Sequence 111, Appl
29	36	87.8	818	17	US-10-433-757-13	Sequence 13, Appl
30	35	85.4	232	15	US-10-424-599-262566	Sequence 262566, A
31	35	85.4	346	16	US-10-791-488A-4	Sequence 4, Appl
32	35	85.4	517	16	US-10-749-714-2	Sequence 2, Appl
33	35	85.4	518	9	US-09-794-927-2	Sequence 2, Appl
34	35	85.4	518	9	US-09-795-847-2	Sequence 2, Appl
35	35	85.4	518	9	US-09-794-743-2	Sequence 2, Appl
36	35	85.4	518	9	US-09-794-748-2	Sequence 2, Appl
37	35	85.4	518	9	US-09-794-925-2	Sequence 2, Appl
38	35	85.4	518	9	US-09-215-450-19	Sequence 19, Appl
39	35	85.4	518	9	US-09-681-442-2	Sequence 2, Appl
40	35	85.4	518	9	US-09-978-295A-196	Sequence 196, Appl
41	35	85.4	518	9	US-09-886-143-2	Sequence 2, Appl
42	35	85.4	518	9	US-09-978-697-196	Sequence 196, Appl
43	35	85.4	518	9	US-09-978-132A-196	Sequence 196, Appl
44	35	85.4	518	9	US-09-999-832A-196	Sequence 196, Appl
45	35	85.4	518	10	US-09-978-189-196	Sequence 196, Appl

## ALIGNMENTS

### RESULT 1

US-09-918-243-86  
; Sequence 86, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 86  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 1-9 of the SCCE protein

### Query Match

Query Match 100.0%; Score 41; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MARSLLPL 9

Db 1 MARSLLPL 9

### RESULT 2

US-09-905-083-86  
; Sequence 86, Application US/09905083  
; Patent No. US20020146708A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; TITLE OF INVENTION: Ovarian Cancer  
; FILE REFERENCE: D6223CIP/C/Div  
; CURRENT APPLICATION NUMBER: US/09/905,083  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 09/502,600  
; PRIOR FILING DATE: 2000-02-11  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 86  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 1-9 of the SCCE protein  
US-09-905-083-86

Query Match 100.0%; Score 41; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
| | | | | | | | |  
Db 1 MARSLLLPL 9

RESULT 3  
US-10-372-521-86  
; Sequence 86, Application US/10372521  
; Publication No. US20030223973A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP2  
; CURRENT APPLICATION NUMBER: US/10/372,521  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 86  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 1-9 of the SCCE protein  
US-10-372-521-86

Query Match 100.0%; Score 41; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
| | | | | | | | |  
Db 1 MARSLLLPL 9

RESULT 4  
US-10-831-075-86  
; Sequence 86, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 86  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 1-9 of the SCCE protein  
US-10-831-075-86

Query Match 100.0%; Score 41; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
| | | | | | | | |  
Db 1 MARSLLLPL 9

RESULT 5  
US-10-262-511-96  
; Sequence 96, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berchs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038



```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/764,762
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,084
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Query Match 100.0%; Score 41; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 8
US-10-071-214-2
; Sequence 2, Application US/10071214
; Publication No. US20030066099A1
; GENERAL INFORMATION:
; APPLICANT: HANSSON, Lennart
; APPLICANT: EGELRUD, Torbjorn
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN
; FILE REFERENCE: HANSSON=3A
; CURRENT APPLICATION NUMBER: US/10/071,214
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,422
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: DK PA 2001 00218
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-071-214-2

Query Match 100.0%; Score 41; DB 14; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9

```

Db 1 MARSLLLPL 9

## RESULT 9

US-10-071-214-48

; Sequence 48, Application US/10071214

; Publication No. US2003006099A1

; GENERAL INFORMATION:

; APPLICANT: HANSSON, Lennart

; APPLICANT: EGEURUD, Torbjorn

; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN

; FILE REFERENCE: HANSSON=3A

; CURRENT APPLICATION NUMBER: US/10/071,214

; CURRENT FILING DATE: 2002-02-11

; PRIOR APPLICATION NUMBER: US 60/267,422

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: DK PA 2001 00218

; PRIOR FILING DATE: 2001-02-09

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 48

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Deduced amino acid sequence from the C-terminal part of SCCE from

; OTHER INFORMATION: homo sapiens.

US-10-071-214-48

Query Match 100.0%; Score 41; DB 14; Length 253;

Best Local Similarity 100.0%; Pred. No. 5.8;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9

|||||

Db 1 MARSLLLPL 9

## RESULT 10

US-10-264-283-90

; Sequence 90, Application US/10264283

; Publication No. US2003014494A1

; GENERAL INFORMATION:

; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

; FILE REFERENCE: 210121.590

; CURRENT APPLICATION NUMBER: US/10/264,283

; CURRENT FILING DATE: 2002-10-02

; NUMBER OF SEQ ID NOS: 111

; SOFTWARE: Corixa Invention Disclosure Database

; SEQ ID NO 90

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-264-283-90

Query Match

Best Local Similarity 100.0%; Score 41; DB 14; Length 253;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9

|||||

Db 1 MARSLLLPL 9

## RESULT 11

US-10-295-027-498

; Sequence 498, Application US/10295027

; Publication No. US20030232350A1

; GENERAL INFORMATION:

; APPLICANT: Afar, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.

; APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

; APPLICANT: Watson, Susan R.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

; FILE REFERENCE: 018501-012500US

; CURRENT APPLICATION NUMBER: US/10/295,027

; CURRENT FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: US 09/663,733

; PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/335,394

; PRIOR FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/332,464

; PRIOR FILING DATE: 2001-11-21

; PRIOR APPLICATION NUMBER: US 60/334,393

; PRIOR FILING DATE: 2001-11-29

; PRIOR APPLICATION NUMBER: US 60/340,376

; PRIOR FILING DATE: 2001-12-14

; PRIOR APPLICATION NUMBER: US 60/347,211

; PRIOR FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US 60/347,349

; PRIOR FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/355,250

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: US 60/356,714

; PRIOR FILING DATE: 2002-02-13

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1386

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 498

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-295-027-498

Query Match 100.0%; Score 41; DB 15; Length 253;

Best Local Similarity 100.0%; Pred. No. 5.8;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9

|||||

Db 1 MARSLLLPL 9

RESULT 12

US-10-173-999-48

; Sequence 48, Application US/10173999

; Publication No. US20040005563A1

; GENERAL INFORMATION:

; APPLICANT: Mack, David H.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions

; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian

; TITLE OF INVENTION: Cancer

; FILE REFERENCE: 018501-002420US

; CURRENT APPLICATION NUMBER: US/10/173,999

; CURRENT FILING DATE: 2002-06-17

; PRIOR APPLICATION NUMBER: US 60/299,234

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: US 60/315,287

; PRIOR FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/372,246

; PRIOR FILING DATE: 2001-04-12  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-173-999-48

Query Match 100.0%; Score 41; DB 15; Length 253;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
|||  
Db 1 MARSLLLPL 9

RESULT 13  
US-10-408-765A-639  
; Sequence 639, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 639  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-639

Query Match 100.0%; Score 41; DB 16; Length 253;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
|||  
Db 1 MARSLLLPL 9

RESULT 14  
US-10-643-795A-95  
; Sequence 95, Application US/10643795A  
; Publication No. US20040241703A1  
; GENERAL INFORMATION:  
; APPLICANT: FREDERIC J. DESAUVAGE  
; APPLICANT: GRETCHEN FRANTZ  
; APPLICANT: KENNETH J. HILLAN  
; APPLICANT: PAUL POLAKIS  
; APPLICANT: ANDREW POLSON  
; APPLICANT: VICTORIA SMITH  
; APPLICANT: SUSAN D. SPENCER  
; APPLICANT: THOMAS D. WU  
; APPLICANT: ZEMIN ZHANG  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5026R1-US  
; CURRENT APPLICATION NUMBER: US/10/643,795A  
; CURRENT FILING DATE: 2003-08-19  
; PRIOR APPLICATION NUMBER: US 60/404,809  
; PRIOR FILING DATE: 2002-08-19

; PRIOR APPLICATION NUMBER: US 60/405,645  
; PRIOR FILING DATE: 2002-08-21  
; PRIOR APPLICATION NUMBER: US 60/413,192  
; PRIOR FILING DATE: 2002-09-23  
; PRIOR APPLICATION NUMBER: US 60/419,008  
; PRIOR FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/426,847  
; PRIOR FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: US 60/484,959  
; PRIOR FILING DATE: 2003-07-02  
; NUMBER OF SEQ ID NOS: 158  
; SEQ ID NO 95  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-643-795A-95

Query Match 100.0%; Score 41; DB 16; Length 253;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
|||  
Db 1 MARSLLLPL 9

RESULT 15  
US-10-948-518-95  
; Sequence 95, Application US/10948518  
; Publication No. US20050064492A1  
; GENERAL INFORMATION:  
; APPLICANT: FREDERIC J. DESAUVAGE  
; APPLICANT: GRETCHEN FRANTZ  
; APPLICANT: KENNETH J. HILLAN  
; APPLICANT: PAUL POLAKIS  
; APPLICANT: ANDREW POLSON  
; APPLICANT: VICTORIA SMITH  
; APPLICANT: SUSAN D. SPENCER  
; APPLICANT: THOMAS D. WU  
; APPLICANT: ZEMIN ZHANG  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5026R1-US  
; CURRENT APPLICATION NUMBER: US/10/948,518  
; CURRENT FILING DATE: 2004-09-22  
; PRIOR APPLICATION NUMBER: US/10/643,795  
; PRIOR FILING DATE: 2003-08-19  
; PRIOR APPLICATION NUMBER: US 60/404,809  
; PRIOR FILING DATE: 2002-08-19  
; PRIOR APPLICATION NUMBER: US 60/405,645  
; PRIOR FILING DATE: 2002-08-21  
; PRIOR APPLICATION NUMBER: US 60/413,192  
; PRIOR APPLICATION NUMBER: US 60/419,008  
; PRIOR FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/426,847  
; PRIOR FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: US 60/484,959  
; PRIOR FILING DATE: 2003-07-02  
; NUMBER OF SEQ ID NOS: 158  
; SEQ ID NO 95  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-948-518-95

Query Match 100.0%; Score 41; DB 17; Length 253;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9  
|||  
Db 1 MARSLLLPL 9

Search completed: July 13, 2005, 18:23:42  
Job time : 65.8889 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:12:23 ; Search time 76.6667 Seconds  
(without alignments)  
45.402 Million cell updates/sec

Title: US-09-905-083a-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	4	Aae08334 Human str
2	42	100.0	9	4	Aae08304 Human str
3	42	100.0	9	8	Adr68891 Human str
4	42	100.0	9	8	Adr68860 Human str
5	42	100.0	97	6	Ada05740 Human NOV
6	42	100.0	97	6	Adh62904 Human NOV
7	42	100.0	144	8	Adi39727 Stratum c
8	42	100.0	144	8	Adi37151 Stratum c
9	42	100.0	181	6	Ada05738 Human NOV
10	42	100.0	181	8	Adh62902 Human NOV
11	42	100.0	198	6	Ada05736 Human NOV
12	42	100.0	198	8	Adh62900 Human NOV
13	42	100.0	224	6	Ada05744 Human NOV
14	42	100.0	224	8	Adh62908 Human NOV
15	42	100.0	225	4	Ada98502 Human Str
16	42	100.0	247	6	Ada05742 Human NOV
17	42	100.0	247	8	Adh62906 Human NOV
18	42	100.0	250	6	Ada05732 Human NOV
19	42	100.0	250	8	Adh62896 Human NOV
20	42	100.0	252	6	Ada05734 Human NOV
21	42	100.0	252	8	Adh62898 Human NOV
22	42	100.0	253	2	Aar67888 Human str
23	42	100.0	253	2	AAW05383 Human any
24	42	100.0	253	5	Abb84421 Human SCC
25	42	100.0	253	5	Abb84406 Human SCC

26	42	100.0	253	5	AAU82740	Aau82740 Amino aci
27	42	100.0	253	6	ABU07440	Abu07440 Protein d
28	42	100.0	253	6	ABU07471	Abu07471 Protein d
29	42	100.0	253	6	ABR58471	ABr58471 Human str
30	42	100.0	253	7	ADB80484	Adb80484 Ovarian c
31	42	100.0	253	7	ADJ68833	Adj68833 Human hea
32	42	100.0	253	7	ADN39180	Adn39180 Cancer/an
33	42	100.0	253	8	ADL06515	Adl06515 Human tum
34	42	100.0	253	8	ADN04182	Adn04182 Antipgori
35	42	100.0	253	8	ADR72880	Adr72880 Human ova
36	42	100.0	257	3	AAb21326	AAb21326 Human HSC
37	39	92.9	226	5	ABB84422	Abb84422 Rat SCCE
38	37	88.1	136	4	ABG23378	Abg23378 Novel hum
39	36	85.7	9	4	AAE08302	Aae08302 Human str
40	36	85.7	9	8	ADR68858	Adr68858 Human str
41	36	85.7	11	5	ABG68630	ABg68630 Stratum c
42	36	85.7	249	5	ABB84423	Abb84423 Murine SC
43	33	78.6	243	5	ABB84419	Abb84419 Bovine SC
44	33	78.6	411	7	ADE40226	Ade40226 Human NOV
45	33	78.6	411	7	ADE40216	Ade40216 Human NOV

## ALIGNMENTS

## RESULT 1

AAE08334  
ID AAE08334 standard; peptide; 9 AA.

XX AC AAE08334;

XX DT 01-NOV-2001 (first entry)

XX DE Human stratum corneum chymotrypsin enzyme peptide #99 (residues 91-99).

XX KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;

XX KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;

XX KW antisense therapy; malignant hyperplasia.

XX OS Homo sapiens.

XX PN WO200159158-A1.

XX PD 16-AUG-2001.

XX PF 07-FEB-2001; 2001WO-US003977.

XX PR 11-FEB-2000; 2000US-00502600.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'Brien TJ;

XX DR WPI; 2001-514676/56.

XX PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX PS Disclosure; Page 124; 127pp; English.

XX CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate is and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 4; Length 9;

```

Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 2
AAE08304
ID AAE08304 standard; peptide; 9 AA.
AC AAE08304;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human stratum corneum chymotrypsin enzyme peptide #69 (residues 91-99).
KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
KW antisense therapy; malignant hyperplasia.
XX
OS Homo sapiens.
XX
PN WO200159158-A1.
XX
PD 16-AUG-2001.
XX
PF 07-FEB-2001; 2001WO-US003977.
XX
PR 11-FEB-2000; 2000US-00502600.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI O'brien TJ;
DR WPI; 2001-514676/56.
XX
PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin
PT enzyme.
XX
PS Claim 25; Page 117; 127pp; English.
XX
CC The invention relates to diagnosing cancer especially ovarian cancer, by
CC screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are
CC considered to be an integral part of tumour growth and metastasis, and
CC therefore, markers indicative of their presence or absence are useful for
CC the diagnosis of cancer. The method is useful for diagnosing cancer,
CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used
CC to treat a cancer selected from ovarian, breast, lung, colon, prostate
CC and other cancers in which SCCE is overexpressed. The present sequence is
CC human SCCE peptide
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 3
ADR68891
ID ADR68891 standard; peptide; 9 AA.
XX
AC ADR68891;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human stratum corneum chymotrypsin enzyme peptide fragment SEQ ID NO:130.

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 4
ADR68860
ID ADR68860 standard; peptide; 9 AA.
XX
AC ADR68860;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human stratum corneum chymotrypsin enzyme peptide fragment SEQ ID NO:99.
KW serine protease; stratum corneum chymotrypsin enzyme; SCCE;
KW immune response; ovarian cancer; lung cancer; prostate cancer;
KW pancreatic cancer; colon cancer.
XX
OS Homo sapiens.
XX
PN WO2004075723-A2.
XX
PD 10-SEP-2004.
XX

```

```

XX serine protease; stratum corneum chymotrypsin enzyme; SCCE;
KW immune response; ovarian cancer; lung cancer; prostate cancer;
KW pancreatic cancer; colon cancer.
XX
OS Homo sapiens.
XX
PN WO2004075723-A2.
XX
PD 10-SEP-2004.
XX
PF 20-FEB-2004; 2004WO-US005134.
XX
PR 21-FEB-2003; 2003US-00372521.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI O'brien TJ, Cannon MJ, Santin A;
XX
DR WPI; 2004-653294/63.
XX
PT Use of stratum corneum chymotrypsin enzyme (SCCE) peptides, for
PT vaccinating an individual against SCCE, and in monitoring and developing
PT immunotherapies for ovarian and other malignancies.
XX
PS Disclosure; SEQ ID NO 130; 117pp; English.
XX
CC The invention relates to a novel method for vaccinating an individual
CC against stratum corneum chymotrypsin enzyme (SCCE), comprising inoculating
CC an individual with a SCCE peptide, which elicits an immune response in
CC the individual. A peptide of the invention acts as a stratum corneum
CC chymotrypsin enzyme inhibitor. The SCCE peptide is useful for vaccinating
CC an individual against SCCE, particularly an individual having, suspected
CC or at risk of getting ovarian, lung, prostate, pancreatic or colon
CC cancer. The oligonucleotide is useful for treating a neoplastic state in
CC an individual, such as ovarian, breast, lung, colon, prostate, or
CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The
CC peptides are also useful in the monitoring and development of
CC immunotherapies for ovarian and other malignancies. The present sequence
CC represents a peptide fragment of serine protease SCCE (stratum corneum
CC chymotrypsin enzyme).
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 4
ADR68860
ID ADR68860 standard; peptide; 9 AA.
XX
AC ADR68860;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human stratum corneum chymotrypsin enzyme peptide fragment SEQ ID NO:99.
KW serine protease; stratum corneum chymotrypsin enzyme; SCCE;
KW immune response; ovarian cancer; lung cancer; prostate cancer;
KW pancreatic cancer; colon cancer.
XX
OS Homo sapiens.
XX
PN WO2004075723-A2.
XX
PD 10-SEP-2004.
XX

```

PF 20-FEB-2004; 2004WO-US005134.  
 PR 21-FEB-2003; 2003US-00372521.  
 XX (UYAR-) UNIV ARKANSAS.  
 PA O'brien TJ, Cannon MJ, Santin A;  
 PI MPI; 2004-653294/63.  
 DR Use of stratum corneum chymotrytic enzyme (SCCE) peptides, for  
 XX vaccinating an individual against SCCE, and in monitoring and developing  
 PT immunotherapies for ovarian and other malignancies.  
 PT Claim 5; SEQ ID NO 99; 117pp; English.  
 XX The invention relates to a novel method for vaccinating an individual  
 CC against stratum corneum chymotrytic enzyme (SCCE), comprising inoculating  
 CC an individual with a SCCE peptide, which elicits an immune response in  
 CC the individual. A peptide of the invention acts as a stratum corneum  
 CC chymotrytic enzyme inhibitor. The SCCE peptide is useful for vaccinating  
 CC an individual against SCCE, particularly an individual having, suspected  
 CC or at risk of getting ovarian, lung, prostate, pancreatic or colon  
 CC cancer. The oligonucleotide is useful for treating a neoplastic state in  
 CC an individual, such as ovarian, breast, lung, colon, prostate, or  
 CC pancreatic cancer, and other cancers in which SCCE is overexpressed. The  
 CC peptides are also useful in the monitoring and development of  
 CC immunotherapies for ovarian and other malignancies. The present sequence  
 CC represents a peptide fragment of serine protease SCCE (stratum corneum  
 CC chymotrytic enzyme).  
 XX Sequence 9 AA;  
 SQ Query Match 100.0%; Score 42; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ORIKASKSF 9  
 Db 1 ORIKASKSF 9  
 RESULT 5  
 ADA05740  
 ID ADA05740 standard; protein; 97 AA.  
 AC ADA05740;  
 XX 06-NOV-2003 (first entry)  
 DE Human NOV18e protein SEQ ID NO:100.  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX Homo sapiens.  
 OS WO2003029424-A2.  
 PN 10-APR-2003.  
 XX 02-OCT-2002; 2002WO-US031373.  
 PF 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 09-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 29-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX (CURA-) CURAGEN CORP.  
 PA Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Szytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton B;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX WPI; 2003-381626/36.  
 DR N-PSDB; ADA05739.  
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX Claim 1; Page 171; 586pp; English.  
 PS The present invention describes NOVX proteins, where X can be 1 to 55  
 XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 97 AA;

Query Match 100.0%; Score 42; DB 6; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.24; 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 1 QRIKASKGF 9  
 |||||  
 Db 72 QRIKASKGF 80

## RESULT 6

ADN62904  
 ID ADN62904 standard; protein; 97 AA.

XX AC ADN62904;

XX DT 01-JUL-2004 (first entry)

XX DE Human NOV18e.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

XX OS Homo sapiens.

XX PN US2004038223-A1.

XX PD 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 09-OCT-2001; 2001US-0327499P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0332666P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENNA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIPI/) DIPIPO V A.  
 PA (EISE/) EISEN A. E.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CEA, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2004-213931/20.

XX DR N-PSDB; ADN62903.

XX PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,

XX PS diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 100; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and



CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.

SQ Sequence 97 AA;

Query Match 100.0%; Score 42; DB 8; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.24; 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
 |||||  
 Db 72 QRIKASKSF 80

RESULT 7  
 ADI39727  
 ID ADI39727 standard; protein; 144 AA.  
 AC ADI39727;  
 XX  
 DT 15-APR-2004 (first entry)  
 XX  
 DE Stratum corneum chymotryptic enzyme (scce) catalytic domain.  
 XX  
 KW Immune T cell; dendritic cell; extracellular serine protease;  
 KW tumour antigen derived gene-14; TADG-14; carcinoma;  
 KW stratum corneum chymotryptic enzyme; scce.

XX Unidentified.  
 OS US6642013-B1.  
 PN  
 XX  
 PD 04-NOV-2003.  
 PF 18-JUL-2000; 2000US-00618259.

PR 21-AUG-1997; 97US-00915659.  
 PR 21-AUG-1998; 98US-00137944.  
 XX  
 PA (UYAR-) UNIV ARKANSAS MEDICAL SCI.  
 PI O'brien TJ, Underwood LJ;  
 XX WPI; 2004-118109/12.

XX Production of activated immune cells or dendritic cells by exposing  
 PT immune cells to tumor antigen derived gene protein fragment consisting of  
 PT amino acid sequences.  
 XX  
 PS Example 1; SEQ ID NO 4; 44pp; English.

CC The present invention relates to novel activated immune T cells or  
 CC dendritic cells directed toward extracellular serine protease termed  
 CC tumour antigen derived gene-14 (TADG-14). The method of the invention  
 CC involves exposing the immune cells to a TADG-14 protein fragment, where  
 CC exposure to the TADG-14 protein fragment activates the immune cells. The  
 CC invention is used for the production of activated immune T cells or  
 CC dendritic cells. The invention allows screening to identify proteases  
 CC overexpressed in carcinoma. The present sequence is stratum corneum  
 CC chymotryptic enzyme (scce) catalytic domain. This sequence is used in the  
 CC invention.

SQ Sequence 144 AA;  
 Query Match 100.0%; Score 42; DB 8; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 0.37; 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
 |||||  
 Db 28 QRIKASKSF 36

RESULT 8  
 ADI37151  
 ID ADI37151 standard; protein; 144 AA.

XX  
 AC ADI37151;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Stratum corneum chymotryptic enzyme (scce) catalytic domain.  
 XX  
 KW Serine protease; tumour antigen derived gene-14; TADG-14;  
 KW neoplastic state; cancer; ovary; breast; lung; colon; prostate;  
 KW chymotryptic enzyme; scce; enzyme.

XX Homo sapiens.  
 OS  
 PN US2003199010-A1.  
 XX  
 PD 23-OCT-2003.

PF 13-JUN-2003; 2003US-00461787.  
 XX  
 PR 21-AUG-1997; 97US-00915659.  
 PR 21-AUG-1998; 98US-00137944.  
 PR 18-JUL-2000; 2000US-00618259.

XX (UYAR-) UNIV ARKANSAS.  
 XX  
 PI O'brien TJ, Underwood LJ;  
 XX WPI; 2004-141550/14.

XX Novel isolated DNA encoding tumor antigen derived gene-14 protein, useful  
 PT for treating neoplastic state (such as ovarian cancer, breast cancer,  
 PT lung cancer, colon cancer, prostate cancer) in an individual.

XX Example 1; SEQ ID NO 4; 46pp; English.

CC The invention relates to extracellular serine protease termed tumour  
 CC antigen derived gene-14 (TADG-14) and its nucleic acid. Composition  
 CC comprising TADG-14 peptide is useful for treating a neoplastic state in  
 CC an individual. The neoplastic state is chosen from ovarian cancer, breast  
 CC cancer, lung cancer, colon cancer, prostate cancer in which TADG-14 is  
 CC overexpressed. The present sequence is Stratum corneum chymotryptic  
 CC enzyme (scce) catalytic domain. This sequence is used in the  
 CC exemplification of the invention.

SQ Sequence 144 AA;

Query Match 100.0%; Score 42; DB 8; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
 |||||  
 Db 28 QRIKASKSF 36

RESULT 9  
 ADA05738  
 ID ADA05738 standard; protein; 181 AA.

XX  
 AC ADA05738;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV18d protein SEQ ID NO:98.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patpurajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zernhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

DR N-PSDB; ADA05737.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX Claim 1; Page 171; 586pp; English.

PS The present invention describes NOVX proteins, where X can be 1 to 55

CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antilipidemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 181 AA;

Query Match 100.0%; Score 42; DB 6; Length 181;

Best Local Similarity 100.0%; Pred. No. 0.47; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0;

QY 1 QRIKASKSF 9

Db 72 QRIKASKSF 80

RESULT 10

ADN62902

ID ADN62902 standard; protein; 181 AA.

XX AC ADN62902;

XX DT 01-JUL-2004 (first entry)

XX DE Human NOV18d.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

XX OS Homo sapiens.

XX PN US2004038223-A1.

XX PD 26-FEB-2004.

XX PF 01-OCT-2002; 2002US-00262511.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 05-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.





PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERR/) ZERRHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATT/) CATTERTON E.  
 PA (JIWW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIPI/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI: 2004-213931/20.  
 DR N-PSDB; ADM62899.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 1; SEQ ID NO 96; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient with  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's disease, Parkinson's Disease, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.

XX Sequence 198 AA;

Query Match 100.0%; Score 42; DB 8; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 0.51;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORIKASKSF 9  
 Db |||||  
 91 ORIKASKSF 99

RESULT 13

ADA05744  
 ID ADA05744 standard; protein; 224 AA.

XX ADA05744;

XX 06-NOV-2003 (first entry)

DE Human NOV189 protein SEQ ID NO:104.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

PN WO2003029424-A2.

PD 10-APR-2003.

PF 02-OCT-2002; 2002WO-US031373.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 12-OCT-2001; 2001US-0328056P.

PR 15-OCT-2001; 2001US-0328499P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 29-OCT-2001; 2001US-0343629P.

PR 01-NOV-2001; 2001US-0349575P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

PA Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI: 2003-381626/36.

DR N-PSDB; ADA05743.

XX

PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.

XX Claim 1; Page 172; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC binding the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virocidic,  
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
 CC and antileptemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX SQ Sequence 224 AA;

Query Match 100.0%; Score 42; DB 6; Length 224;  
 Best Local Similarity 100.0%; Pred No. 0.58;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QRIKASKSF 9

Db |||||  
 77 QRIKASKSF 85

RESULT 14

ADN62908

ID ADN62908 standard; protein; 224 AA.

XX AC ADN62908;

XX DT 01-JUL-2004 (first entry)

XX DE Human NOV18g.

XX KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

XX KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;

XX KW Alzheimer's disease; Parkinson's disease; immune disorder;

XX KW haematopoietic disorder; dyslipidaemia; metabolic syndrome X;

XX KW wasting disorder.

XX OS Homo sapiens.

XX XX

PN US2004038223-A1.

XX

PD 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 17-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

XX 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.

PA (JUJJ/) JU J.

PA (LILL/) LI L.

PA (GUOX/) GUO X.

PA (PATT/) PATTURAJAN M.

PA (SPYT/) SPYTEK K A.

PA (EDIN/) EDINGER S R.

PA (ELLE/) ELLERMAN K.

PA (MALY/) MALYANKAR U M.

PA (ORTT/) ORT T.

PA (GORM/) GORMAN L.

PA (ZERH/) ZERHUSEN B D.

PA (ANDE/) ANDERSON D W.

PA (ZHON/) ZHONG M.

PA (CATT/) CATTERTON E.

PA (JIWW/) JI W.

PA (MILL/) MILLER C E.

PA (RAST/) RASTELLI L.

PA (STON/) STONE D J.

PA (PENA/) PENNA C E A.

PA (SHEN/) SHENOV S G.

PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.

PA (LEAC/) LEACH M D.

PA (AGEE/) AGEE M L.

PA (BERG/) BERGHS C.

PA (DIP/) DIPIPO V A.

PA (EISE/) EISEN A.

PA (GANG/) GANGOLLI E A.

PA (RIEG/) RIEGER D K.

PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malvankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX WPI; 2004-213931/20.  
 DR N-PSDB; ADN62907.  
 XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 PT Claim 1; SEQ ID NO 104; 395pp; English.  
 XX The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents the amino acid sequence of a human NOVX protein.  
 XX Sequence 224 AA;  
 SQ

Query Match 100.0%; Score 42; DB 8; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 0.58;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
 Db 77 QRIKASKSF 85  
 |||||  
 |||||

RESULT 15  
 AAB98502  
 ID AAB98502 standard; protein; 225 AA.  
 XX AC AAB98502;  
 XX 03-AUG-2001 (first entry)  
 XX Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.  
 KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;  
 KW tumour antigen-derived gene 15; serine protease;  
 KW Stratum Corneum Chymotryptic Enzyme; SCCE.  
 XX Homo sapiens.  
 OS WO200129056-A1.  
 PN 26-APR-2001.  
 XX 20-OCT-2000; 2000WO-US029095.

PR 20-OCT-1999; 99US-00421213.  
 XX (UYAR-) UNIV ARKANSAS.  
 XX O'brien TJ, Tanimoto H;  
 XX WPI; 2001-381031/40.  
 XX Novel extracellular serine protease, termed tumor antigen-derived gene 15  
 PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,  
 PT treatment, prevention of cancer, particularly breast, ovarian cancer.  
 XX Example 10; Fig 1; 130pp; English.  
 XX The present invention relates to human tumour antigen-derived gene 15  
 CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500). TADG-  
 CC 15 is an extracellular serine protease. It was found that TADG-15 is over-  
 CC expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20  
 CC residues that lack TADG-15 protease activity are useful for vaccinating  
 CC an individual against TADG-15, having, suspected of having or at risk of  
 CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic  
 CC or therapeutic target in cancer. The present sequence was used in a  
 CC sequence homology alignment with the catalytic domain of TADG-15  
 XX Sequence 225 AA;  
 SQ

Query Match 100.0%; Score 42; DB 4; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 0.58;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
 Db 63 QRIKASKSF 71  
 |||||  
 |||||

Search completed: July 13, 2005, 17:19:57  
 Job time : 77.6667 secs

**This page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:54:03 ; Search time 13.5556 Seconds  
(without alignments)  
63.882 Million cell updates/sec

Title: US-09-905-083A-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	253	2 A53968	serine proteinase
2	34	81.0	373	2 D89004	protein T24A6.11 (
3	33	78.6	427	2 A49518	kallistatin precu
4	32	76.2	895	2 D72571	probable DNA-direc
5	31	73.8	248	2 A40625	tetrachloro-p-hydr
6	31	73.8	298	2 S32368	beta-SNAP protein
7	31	73.8	302	2 T50147	mak16-like protein
8	31	73.8	362	2 F81409	probable phosphate
9	30	71.4	142	2 C39610	BET1 protein - yea
10	30	71.4	321	2 C71463	probable p-loop AT
11	30	71.4	331	2 H97051	probable dehydroge
12	30	71.4	427	2 T48159	hypothetical prote
13	30	71.4	517	2 T40129	hypothetical prote
14	30	71.4	627	2 C86850	neutral endopeptid
15	30	71.4	627	2 A47098	lactococcal endope
16	30	71.4	627	2 F53290	endopeptidase PepO
17	30	71.4	875	2 T19678	hypothetical prote
18	30	71.4	985	2 D82776	pyruvate dehydroge
19	30	71.4	1102	2 H84545	probable ubiquitin
20	30	71.4	3660	1 S02041	dystrophin, muscle
21	29	69.0	65	2 AF2269	hypothetical prote
22	29	69.0	141	1 B46322	16k protein - toba
23	29	69.0	172	2 H81251	NADH2 dehydrogenas
24	29	69.0	182	2 C64176	hypothetical prote
25	29	69.0	203	2 C90678	thiogalactoside ac
26	29	69.0	203	2 G85528	thiogalactoside ac
27	29	69.0	268	2 E84339	quinolinate phosph
28	29	69.0	349	2 A25268	WD repeat protein
29	29	69.0	348	2 T22587	hypothetical prote

30	29	69.0	463	2 G97300	aspartyl/asparagin
31	29	69.0	484	2 JC5282	phosphogluconate d
32	29	69.0	526	1 G71081	probable helicase
33	29	69.0	563	2 B82883	hypothetical prote
34	29	69.0	584	2 T48273	hypothetical prote
35	29	69.0	585	2 I37216	calicin - human (f
36	29	69.0	669	2 AB2028	hypothetical prote
37	29	69.0	691	2 T46476	hypothetical prote
38	29	69.0	717	2 AC1419	DNA topoisomerase
39	29	69.0	738	2 E86294	hypothetical prote
40	29	69.0	743	2 C86168	hypothetical prote
41	29	69.0	759	2 S67164	probable membrane
42	29	69.0	940	2 A40985	projectin - fruit
43	29	69.0	1018	2 T15297	hypothetical prote
44	29	69.0	1742	2 S24600	projectin - fruit
45	29	69.0	6658	2 T13931	projectin - fruit

ALIGNMENTS

RESULT 1

A53968

serine proteinase SCCE precursor - human

N;Alternate names: stratum corneum chymotryptic enzyme

C;Species: Homo sapiens (man)

C;Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C;Accession: A53968

R;Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A;Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A;Reference number: A53968; MUID:94308225; PMID:8034709

A;Accession: A53968

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-253 <HAN>

A;Cross-references: UNIPROT:P49862; GB:L33404; NID:g521214; PIDN:AAC37551.1; PID:g532504

C;Genetics: A;Gene: GDB:PRSS6; SCCE

A;Cross-references: GDB:377730

A;Map position: 7q35-7q35

C;Superfamily: trypsin; trypsin homology

F;30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 42; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.096;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

RESULT 2

D89004

protein T24A6.11 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C;Accession: D89004

R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_eleg

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: D89004

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-373 <STO>

A;Cross-references: UNIPROT:O61942; GB:chr V; PIDN:AAC17798.1; PID:g3168946; GSPDB:GN000;

A;Note: contains similarity to C4-type zinc fingers

C;Genetics: A;Gene: T24A6.11

A:Map position: 5

Query Match 81.0%; Score 34; DB 2; Length 373;  
Best Local Similarity 77.8%; Pred. No. 9.3; Mismatches 1; Indels 0; Gaps 0;  
Matches 7; Conservative 1

Qy 1 QRIKASKSF 9  
|:|:|:|:  
Db 9 QEIKSKSF 17

#### RESULT 3

A49518  
kallistatin precursor - human  
N:Alternate names: proteinase inhibitor 4; tissue kallikrein inhibitor  
C:Species: Homo sapiens (man)  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: A49518; A55560; A45093  
R:Chai, K.X.; Chen, L.M.; Chao, J.; Chao, L.  
J. Biol. Chem. 268, 24498-24505, 1993  
A:Title: Kallistatin: a novel human serine proteinase inhibitor. Molecular cloning, tissue  
A:Reference number: A49518; MUID:94043294; PMID:8227002  
A:Accession: A49518  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-427 <CHA>  
A:Cross-references: UNIPROT:P29622; GB:L19684; NID:9425145; PIDN:AAA59454.1; PID:g425146  
R:Chai, K.X.; Ward, D.C.; Chao, J.; Chao, L.  
Genomics 23, 370-378, 1994  
A:Title: Molecular cloning, sequence analysis, and chromosomal localization of the human  
A:Reference number: A55560; MUID:95137583; PMID:7835886  
A:Accession: A55560  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-427 <CH2>  
A:Cross-references: GB:L28101; NID:9609489; PIDN:AAC41706.1; PID:g619783  
R:Zhou, G.X.; Chao, L.; Chao, J.  
J. Biol. Chem. 267, 25873-25880, 1992  
A:Title: Kallistatin: a novel human tissue kallikrein inhibitor. Purification, character  
A:Reference number: A45093; MUID:93100304; PMID:1334488  
A:Accession: A45093  
A:Molecule type: protein  
A:Residues: 389-403 <ZHO>  
A:Cross-references: PIDN:AAB24557.1; PID:g261993  
A:Note: sequence extracted from NCBI backbone (NCBIP:120847)  
C:Genetics:  
A:Gene: GDB:P14; KIST  
A:Cross-references: GDB:266537; OMIM:147935  
A:Map position: 14q31-14q32.1  
C:Superfamily: Serpin  
C:Keywords: blocked amino end; glycoprotein; pyroglutamic acid; serine proteinase inhibi  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-427/Product: kallistatin #status predicted <MAT>  
F:21/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:33,108,157,238/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:388/Inhibitory site: Phe (tissue kallikrein) #status experimental

Query Match 78.6%; Score 33; DB 2; Length 427;  
Best Local Similarity 66.7%; Pred. No. 18; Mismatches 3; Indels 0; Gaps 0;  
Matches 6; Conservative 3

Qy 1 QRIKASKSF 9  
|:|:|:|:  
Db 355 QKLEASKSF 363

#### RESULT 4

D72571  
probable DNA-directed RNA polymerase subunit A' APE1853 - Aetopyrum pernix (strain K1)  
C:Species: Aetopyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: D72571  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: D72571  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-895 <KAW>  
A:Cross-references: UNIPROT:Q9YAU1; DDBJ:AP0000062; NID:g5105244; PIDN:BAA80857.1; PID:g51  
C:Genetics:  
A:Gene: APE1853  
C:Superfamily: Halobacterium DNA-directed RNA polymerase chain A

Query Match 76.2%; Score 32; DB 2; Length 895;  
Best Local Similarity 75.0%; Pred. No. 64; Mismatches 2; Indels 0; Gaps 0;  
Matches 6; Conservative 2

Qy 1 QRIKASKS 8  
|:|:|:|:  
Db 885 QRVKASKA 892

#### RESULT 5

A40625  
tetrachloro-p-hydroquinone reductive dehalogenase (EC 1.-.-.-) - Flavobacterium sp.  
N:Alternate names: Glutathione S-transferase homolog; TeCH reductive dehalogenase  
C:Species: Flavobacterium sp.  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Oct-1999  
C:Accession: A40625  
R:Orser, C.S.; Dutton, J.; Lange, C.; Jablonski, P.; Xun, L.; Hargis, M.  
J. Bacteriol. 175, 2640-2644, 1993  
A:Title: Characterization of a Flavobacterium Glutathione S-transferase gene involved rec  
A:Reference number: A40625; MUID:93239690; PMID:8478329  
A:Accession: A40625  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-248 <ORS>  
A:Cross-references: GB:M98559; NID:g148689; PIDN:AAA24921.1; PID:g148690  
A:Experimental source: ATCC 39723  
A:Note: sequence extracted from NCBI backbone (NCBIN:130432, NCBIP:130433)  
C:Keywords: oxidoreductase

Query Match 73.8%; Score 31; DB 2; Length 248;  
Best Local Similarity 66.7%; Pred. No. 30; Mismatches 2; Indels 1; Gaps 0;  
Matches 6; Conservative 2

Qy 1 QRIKASKSF 9  
|:|:|:|:  
Db 227 QRMKARRSF 235

#### RESULT 6

S32368  
beta-SNAP protein - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S32368  
R:Whiteheart, S.W.; Griff, I.C.; Brunner, M.; Clary, D.O.; Mayer, T.; Buhrow, S.A.; Rothn  
Nature 362, 353-355, 1993  
A:Title: SNAP family of NSF attachment proteins includes a brain-specific isoform.  
A:Reference number: S32367; MUID:93205122; PMID:8455721  
A:Accession: S32368  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-298 <WHI>  
A:Cross-references: UNIPROT:P81126  
C:Superfamily: alpha-soluble NSF attachment protein

Query Match 73.8%; Score 31; DB 2; Length 298;  
Best Local Similarity 66.7%; Pred. No. 36; Mismatches 2; Indels 1; Gaps 0;  
Matches 6; Conservative 2

Qy 1 QRIKASKSF 9  
Db 19 KRVKASHSF 27

RESULT 7  
T50147  
mak16-like protein [imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 12-Jul-2004  
A:Accession: T50147  
R:Hamlin, N.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, November 1999  
A:Reference number: 225043  
A:Accession: T50147  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-302 <HAM>  
A:Cross-references: UNIPROT:Q9UTE6; EMBL:AL132798; PIDN:CA860698.1; GSPDB:GN00066; SPDB:  
A:Experimental source: strain 972h(-); cosmid c222  
C:Genetics:  
A:Gene: SPDB:SPAC222.06  
A:Map position: 1  
A:Introns: 21/2  
C:Superfamily: ribosomal biogenesis protein MAK16

Query Match 73.8%; Score 31; DB 2; Length 302;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
Db 78 QRIKLKSNY 86

RESULT 8  
F81409  
probable phosphate transport system permease protein Cj0615 [imported] - Campylobacter j  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: F81409  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: F81409  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-362 <PAR>  
A:Cross-references: UNIPROT:Q9PHQ2; GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CA87525  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: psta; Cj0615

Query Match 73.8%; Score 31; DB 2; Length 362;  
Best Local Similarity 77.8%; Pred. No. 43;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
Db 7 KROKASKSF 15

RESULT 9  
C39610  
Brl1 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YIA4; protein YIL004c; SLX12 protein  
C:Species: Saccharomyces cerevisiae  
C:Date: 13-Sep-1991 #sequence\_revision 13-Sep-1991 #text\_change 09-Jul-2004  
A:Accession: C39610; S48453; S58677  
R:Daecher, C.; Ossig, R.; Gallwitz, D.; Schmitt, H.D.  
Mol. Cell. Biol. 11, 872-885, 1991

A:Title: Identification and structure of four yeast genes (SLX) that are able to suppress  
A:Reference number: A39610; MUID:91117254; PMID:1990290  
A:Accession: C39610  
A:Molecule type: DNA  
A:Residues: 1-142 <DAS>  
A:Cross-references: UNIPROT:P22804; EMBL:X54237; NID:g4483; PIDN:CAA38143.1; PID:g4484  
R:Rowley, N.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: S48442  
A:Accession: S48453  
A:Molecule type: DNA  
A:Residues: 1-142 <ROW>  
A:Cross-references: GB:Z47047; EMBL:Z38113; NID:g603997; PID:g763342; MIPS:YIL004c  
R:Voas, H.; Tanames, J.; Teodoru, C.; Valencia, A.; Sensen, C.; Wiemann, S.; Schwager, C.  
Yeast 11, 61-78, 1995  
A:Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome 1  
A:Reference number: S50795; MUID:95282515; PMID:7762303  
A:Accession: S58677  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-142 <VOS>  
A:Cross-references: EMBL:X79743  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994  
C:Genetics:  
A:Gene: SGD:BET1; SLX12  
A:Cross-references: SGD:S0001266; MIPS:YIL004c  
A:Map position: 9L  
A:Introns: 4/2  
C:Keywords: Transmembrane protein  
F:126-142/Domain: transmembrane #status predicted <TMM>

Query Match 71.4%; Score 30; DB 2; Length 142;  
Best Local Similarity 87.5%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKS 8  
Db 67 QRIKALKS 74

RESULT 10  
C71463  
probable p-loop ATPase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C:Accession: C71463  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia traci  
A:Reference number: A71570; MUID:99000809; PMID:9784136  
A:Accession: C71463  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-321 <ARN>  
A:Cross-references: UNIPROT:O84847; GB:AE001357; GB:AE001273; NID:g3329311; PIDN:AAC6843;  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: mesJ

Query Match 71.4%; Score 30; DB 2; Length 321;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
Db 277 ERLVASKSF 285

RESULT 11  
H97051  
probable dehydrogenase, YULF B. subtilis ortholog [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C;Accession: H97051  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum ATCC824  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: H97051  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-331 <KUR>  
A;Cross-references: UNIPROT:Q97JP2; GB:AE001437; PIDN:AAK79203.1; PID:gl5024156; GSPDB:CAC1231  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
C;Superfamily: Escherichia coli ygjR protein

Query Match 71.4%; Score 30; DB 2; Length 331;  
Best Local Similarity 75.0%; Pred. No. 67;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKS 8  
|:|:|:|:|  
Db 303 QRLKVS 310

RESULT 12  
T48159  
hypothetical protein T1008.80 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T48159  
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Lemcke, K.; Mayer, K. submitted to the Protein Sequence Database, March 2000  
A;Reference number: 224486  
A;Accession: T48159  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-427 <BEV>  
A;Cross-references: UNIPROT:Q9M035; EMBL:AL161746  
A;Experimental source: cultivar Columbia; BAC clone T1008  
C;Genetics:  
A;Map position: 5  
A;Introns: 317/1  
A;Note: T1008.80

Query Match 71.4%; Score 30; DB 2; Length 427;  
Best Local Similarity 66.7%; Pred. No. 87;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
|:|:|:|:|  
Db 140 RRVKASVSF 148

RESULT 13  
T40129  
hypothetical protein SPBC2F12.12c - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T40129  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M. submitted to the EMBL Data Library, June 1997  
A;Reference number: 221907  
A;Accession: T40129  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-517 <WOO>  
A;Cross-references: UNIPROT:O14342; EMBL:Z97211; PIDN:CAB10159.1; GSPDB:GN00067; SPDB:SPBC2F12.12c  
A;Experimental source: strain 972h; cosmid c2F12  
C;Genetics:  
A;Gene: SPDB:SPBC2F12.12c  
A;Map position: 2  
A;Introns: 42/3; 126/3; 157/3; 308/3

C;Accession: H97051  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum ATCC824  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: H97051  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-331 <KUR>  
A;Cross-references: UNIPROT:Q97JP2; GB:AE001437; PIDN:AAK79203.1; PID:gl5024156; GSPDB:CAC1231  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
C;Superfamily: Escherichia coli ygjR protein

Query Match 71.4%; Score 30; DB 2; Length 331;  
Best Local Similarity 75.0%; Pred. No. 67;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
|:|:|:|:|  
Db 187 QRILAPKSF 195

RESULT 14  
C86850  
neutral endopeptidase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: C86850  
R;Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp  
A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: C86850  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-627 <STO>  
A;Cross-references: UNIPROT:Q07744; GB:AE005176; PID:gl2724829; PIDN:AAK05901.1; GSPDB:GCA1231  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: pepO

Query Match 71.4%; Score 30; DB 2; Length 627;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RIKASKSF 9  
|:|:|:|:|  
Db 570 RMKASKEF 577

RESULT 15  
A47098  
lactococcal endopeptidase (EC 3.4.21.1) - Lactococcus lactis subsp. cremoris (strain P8-2-  
C;Species: Lactococcus lactis subsp. cremoris  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: A47098  
R;Mierau, I.; Tan, P.S.T.; Haandrikman, A.J.; Kok, J.; Leenhouts, K.J.; Konings, W.N.; Ve J. Bacteriol. 175, 2087-2096, 1993  
A;Title: Cloning and sequencing of the gene for a lactococcal endopeptidase, an enzyme w  
A;Reference number: A47098; MUID:93209954; PMID:8458851  
A;Accession: A47098  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-627 <MIE>  
A;Cross-references: UNIPROT:Q09145; GB:L04938; NID:g293013; PIDN:AAA25204.1; PID:g293015  
C;Genetics:  
A;Gene: pepO  
C;Keywords: hydrolase; oligopeptide transport

Query Match 71.4%; Score 30; DB 2; Length 627;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RIKASKSF 9  
|:|:|:|:|  
Db 570 RMKASKEF 577

Search completed: July 13, 2005, 17:31:34  
Job time : 15.5556 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 16:15:23 ; Search time 62.2222 Seconds  
(without alignments)  
74.069 Million cell updates/sec

Title: US-09-905-083A-99  
Perfect score: 42  
Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	42	100.0	181	2	Q8NFV7	Q8NFV7 homo sapien
2	42	100.0	253	1	KLK7_HUMAN	P49862 homo sapien
3	42	100.0	253	2	Q8NSN9	Q8NSN9 homo sapien
4	36	85.7	234	2	Q9R048	Q9R048 mus musculus
5	36	85.7	249	2	Q91VE3	Q91VE3 m thymopain
6	34	81.0	107	1	Y197_PASMU	Q9CJU4 pasteurella
7	34	81.0	294	2	Q88Q37	Q88Q37 pseudomonas
8	34	81.0	683	2	Q82XB0	Q82XB0 nitrosomonas
9	33	78.6	310	2	Q678H2	Q678H2 lymphocysti
10	33	78.6	314	1	RLA0_THEAC	P57692 thermoplasma
11	33	78.6	427	1	KAIN_HUMAN	P29622 homo sapien
12	33	78.6	445	2	Q86TE9	Q86TE9 homo sapien
13	33	78.6	671	2	Q8EUP9	Q8EUP9 mycoplasma
14	32	76.2	161	1	RL10_MYCPE	Q8AVJ0 mycoplasma
15	32	76.2	240	2	Q6FDR5	Q6FDR5 acinetobact
16	32	76.2	274	2	Q9U0Z3	Q9U0Z3 leishmania
17	32	76.2	458	2	Q7RKQ8	Q7RKQ8 plasmodium
18	32	76.2	467	2	Q8IBN4	Q8IBN4 plasmodium
19	32	76.2	474	2	Q7VRX5	Q7VRX5 candidatus
20	32	76.2	490	2	Q8SAQ6	Q8SAQ6 chlamydomon
21	32	76.2	684	2	Q816K2	Q816K2 holotrichia
22	32	76.2	761	2	Q6D7X6	Q6D7X6 erwinia car
23	32	76.2	772	2	Q62HP6	Q62HP6 burkholderi
24	32	76.2	772	2	Q63W23	Q63W23 burkholderi
25	32	76.2	793	2	Q7QXY1	Q7QXY1 giardia lam
26	32	76.2	895	2	Q9YAU1	Q9YAU1 aeropyrum p
27	31	73.8	155	2	Q6DGM1	Q6DGM1 brachydanio
28	31	73.8	176	2	Q6AXN0	Q6AXN0 rattus norv
29	31	73.8	176	2	Q6P3B2	Q6P3B2 mus musculus
30	31	73.8	177	2	Q8TB05	Q8TB05 homo sapien
31	31	73.8	212	2	Q71MF6	Q71MF6 homo sapien

32	31	73.8	235	2	Q6PBI0	Q6PBI0 acinetobact
33	31	73.8	247	1	PCPC_SPHCR	Q03520 sphingobium
34	31	73.8	248	2	Q937X0	Q937X0 sphingomona
35	31	73.8	254	2	Q975X7	Q975X7 sulfolobus
36	31	73.8	276	2	Q8N8N1	Q8N8N1 homo sapien
37	31	73.8	298	1	SNAB_BOVIN	P81126 bos taurus
38	31	73.8	298	1	SNAB_HUMAN	Q9H115 homo sapien
39	31	73.8	298	1	SNAB_MOUSE	P28663 mus musculus
40	31	73.8	302	2	Q9UT56	Q9UTE6 schizosacch
41	31	73.8	362	2	Q9PHQ2	Q9PHQ2 campylobact
42	31	73.8	405	2	Q64WE2	Q64WE2 bacteroides
43	31	73.8	433	2	Q6MAE4	Q6MAE4 paracitlamyd
44	31	73.8	443	2	Q8TA05	Q8TA05 acrasia ros
45	31	73.8	895	2	Q8P4T1	Q8P4T1 xanthomonas

#### ALIGNMENTS

RESULT 1  
Q8NFV7 ID Q8NFV7 PRELIMINARY; PRT; 181 AA.  
AC Q8NFV7;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Kallikrein 7 short variant protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovarian carcinoma;  
RX MEDLINE=22623266; PubMed=12738725;  
RA Dong Y., Kauschal A., Brattsand M., Nicklin J., Clements J.A.;  
RT "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer  
produces novel variants with potential as cancer biomarkers.";  
RL Clin. Cancer Res. 9:1710-1720(2003).  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AF411215; AAN03663.1; -;  
DR HSSP; P00760; 1EZX.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR003003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYSPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ HYDROLASE 181 AA; 19887 MW; 86A28A03B80C2D78 CRC64;  
Query Match 100.0%; Score 42; DB 2; Length 181;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QRIKASKSF 9  
DB 19 QRIKASKSF 27  
RESULT 2  
KLK7\_HUMAN ID KLK7\_HUMAN STANDARD; PRT; 253 AA.  
AC P49862;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic  
 GN enzyme) (hSCCE).  
 OS Name=KLK7; Synonyms=PRSS6, SCCE;  
 OC Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.  
 RC TISSUE=Skin;  
 RX MEDLINE=94308225; PubMed=8034709;  
 RA Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A.,  
 RA Egelrud T.;  
 RT "Cloning, expression, and characterization of stratum corneum  
 RT chymotryptic enzyme. A skin-specific human serine proteinase.";  
 RL J. Biol. Chem. 269:19420-19426(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Keratinocytes;  
 RA Yousef G.M., Scorilas A., Diamandis E.P.;  
 RT "Molecular characterization, mapping and tissue expression of the  
 RT human stratum corneum chymotryptic enzyme gene.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepel B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
 RA Wallbrandt P., Egelrud T.;  
 RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in  
 RT mice; a model for chronic itchy dermatitis.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95314630; PubMed=7794273;  
 RA Skyt A., Stroemqvist M., Egelrud T.;  
 RT "Primary substrate specificity of recombinant human stratum corneum  
 RT chymotryptic enzyme.";  
 RL Biochem. Biophys. Res. Commun. 211:586-589(1995).  
 CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive  
 CC structures in the cornified layer of the skin in the continuous  
 CC shedding of cells from the skin surface. Specific for amino acid  
 CC residues with aromatic side chains in the P1 position. SCCE  
 CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-  
 CC Phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the  
 CC activation of precursors to inflammatory cytokines.  
 CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is  
 CC expressed by keratinocytes in the epidermis. Very low levels are  
 CC also seen in the brain and kidney.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [licenses@isb-sib.ch](mailto:licenses@isb-sib.ch)).  
 CC -----  
 CC EMBL; L33404; AAC37551.1; -;  
 CC EMBL; AF166330; AAD49718.1; -;  
 CC EMBL; AF243527; AAG33360.1; -;  
 CC EMBL; AF332583; AAG69624.1; -;  
 CC PIR; A53968; A53968.  
 CC HSP; P00760; IEXZ.  
 CC MEROPS; S01.300; -.

DR Genew; HGNC:6368; KLK7.  
 DR H-InvDB; HIX0015373; -;  
 DR MIM; 604438; -;  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR GO; GO:0008444; P:epidermal differentiation; TAS.  
 DR InterPro; IPR009003; Pept Ser Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SMO0020; Tryp\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;  
 KW Signal; Zymogen.  
 FT SIGNAL 1 22  
 FT PROPEP 23 29  
 FT CHAIN 30 253  
 FT ACT\_SITE 70 70  
 FT ACT\_SITE 112 112  
 FT ACT\_SITE 205 205  
 FT DISULFID 36 137  
 FT DISULFID 55 71  
 FT DISULFID ? 239  
 FT DISULFID 144 211  
 FT DISULFID 176 190  
 FT DISULFID 201 226  
 FT CARBOHYD 246 246  
 SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;  
 Query Match 100.0%; Score 42; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.41;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QRIKASKSF 9  
 Db 91 QRIKASKSF 99  
 RESULT 3  
 Q8NSN9 ID Q8NSN9 PRELIMINARY; PRT; 253 AA.  
 AC Q8NSN9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Stratum corneum chymotryptic enzyme, preproprotein.  
 GN Name=KLK7;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Maman A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfield G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.,  
 RA Krzywinski M.I., Skaleka U., Smallos D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; BC032005; AAH32005.1; -;  
 DR HSSP; P00760; 1EZX.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWNW\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A68 CRC64;  
  
 Query Match 100.0%; Score 42; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.41;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 QRIKASKSF 9  
 Db 91 QRIKASKSF 99  
  
 RESULT 4  
 Q9R048 PRELIMINARY; PRT; 234 AA.  
 AC Q9R048;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
 DE Stratum corneum chymotryptic enzyme (fragment).  
 GN Name=Klk7;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL;  
 RA Backman A., Stranden P., Bratteand M., Hansson L., Egelrud T.;  
 RT "Molecular cloning and tissue expression of the murine analog to human  
 stratum corneum chymotryptic enzyme."; J. Invest. Dermatol. 113:152-155(1999).  
 RL J. Invest. Dermatol. 113:152-155(1999).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; AF124299; AAF01139.1; -;  
 DR HSSP; P00760; 1EZX.  
 DR MEROPS; S01.300; -;  
 DR MGD; MGI:1346336; Klk7.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWNW\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 234 234  
 SQ SEQUENCE 234 AA; 25220 MW; BEBE57D6C325B85F CRC64;  
  
 Query Match 85.7%; Score 36; DB 2; Length 234;  
 Best Local Similarity 77.8%; Pred. No. 9.1;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 QRIKASKSF 9  
 Db 87 QRIKASKSF 95  
  
 RESULT 5  
 Q9IVE3 PRELIMINARY; PRT; 249 AA.  
 ID Q9IVE3  
 AC Q9IVE3;  
 DT 01-DEC-2001 (TREMELrel. 19, Created)  
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
 DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)  
 DE Thymopsin (Kallikrein 7) (Stratum corneum chymotryptic enzyme) (Mus  
 musculus 0 day neonate head cDNA, RIKEN full-length enriched library,  
 clone:4833444E03 product:kallikrein 7 (chymotryptic, stratum corneum),  
 full insert sequence) (Mus musculus 10 days neonate head cDNA, RIKEN  
 full-length enriched library, clone:553040G09 product:kallikrein 7  
 (chymotryptic, stratum corneum), full insert sequence).  
 GN Name=Klk7;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymus;  
 RA Yamaguchi N.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 Krzywanski M.I., Skalska U., Smalls D.B., Schnerch A., Schein J.E.,  
 Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ;  
 RA Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,  
 RA Wallbrandt P., Egelrud T.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44 (1999).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690 (2001).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA The FANTOM Consortium;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630 (2000).  
RN [9]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771 (2000).  
RN [10]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,  
RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Horii F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido Y., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,  
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.,  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

-!- SIMILARITY: Belongs to peptidase family S1.  
CC EMBL; AB008371; BAB55604.1; -  
DR EMBL; BC027823; AAH27823.1; -  
DR EMBL; AF339930; AAK69652.1; -  
DR EMBL; AK029477; BAC26467.1; -  
DR EMBL; AK077406; BAC36787.1; -  
DR HSSP; P00760; 1EZX.  
DR MEROPS; S01.300; -  
DR MGD; MGI-1345336; Kik7  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM0020; TRYD SPC; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR HydroLase; Protease; Serine protease.  
SQ SEQUENCE 249 AA; 27257 MW; 0D4E380F12D14F87 CRC64;  
KW  
Query Match 85.7%; Score 36; DB 2; Length 249;  
Best Local Similarity 77.8%; Pred. No. 9.8;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QRIKASKSF 9  
Db 87 QKIKATKSF 95  
RESULT 6  
YI97\_PASMU STANDARD; PRT; 107 AA.  
ID AC Q9CJU4;  
AD 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Hypothetical lipoprotein PM1897 precursor.  
GN OrderedLocusNames=PM1897;  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
ON NCBI\_TaxID=747;  
RX NCBI\_SEQUENCE FROM N.A.  
RC STRAIN=Pm70;  
EX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
(Potential).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC  
CC EMBL; AE006227; AAK03981.1; -  
DR InterPro; IPR000437; Prok\_LipoProt\_S.  
DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
KW Complete proteome; Hypothetical protein; Lipoprotein; Membrane;  
KW Palmitate; Signal.  
FT SIGNAL 1 17 Potential.  
FT CHAIN 18 107 Hypothetical protein PM1897.  
FT LIPID 18 18 N-palmitoyl cysteine (Potential).  
FT LIPID 18 18 S-diacylglycerol cysteine (Potential).  
SQ SEQUENCE 107 AA; 12214 MW; 920C61E14A10F405 CRC64;  
KW  
Query Match 81.0%; Score 34; DB 1; Length 107;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



```

QY 1 QRIKASKSF 9
Db 27 QRIEAGKSF 35

RESULT 7
Q88Q37 PRELIMINARY; PRT; 294 AA.
AC Q88Q37;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcriptional regulator AmpR, putative.
GN OrderedLocustNames=PP0661;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,
RA Moazzes A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Dueterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
CC -!- SIMILARITY: Contains 1 HTH lySR-type DNA-binding domain.
DR EMBL; AE016776; AAN66286.1; --
DR TIGR; PP0661; --
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000847; HTH_LySR.
DR InterPro; IPR005119; LySR_subst.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LySR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LYSR; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 294 AA; 32995 MW; 33695F24D26560EF CRC64;

Query Match 81.0%; Score 34; DB 2; Length 294;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 259 QRIEASESF 267

RESULT 8
Q82XB0 PRELIMINARY; PRT; 683 AA.
AC Q82XB0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sensory transduction histidine kinases.
GN OrderedLocustNames=NE0377;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]

SEQUENCE FROM N.A.
STRAIN=ATCC 19718 / IFO 14298;
MEDLINE=22586410; PubMed=12700255;
DOI=10.1128/JB.185.9.2759-2773.2003;
Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
Arciero D.M., Holmes N.G., Whittaker M.M., Arp D.J.;
"Complete genome sequence of the ammonia-oxidizing bacterium and
obligate chemolithoautotroph Nitrosomonas europaea.";
J. Bacteriol. 185:2759-2773(2003).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; BX321857; CAD84288.1; --
DR HSSP; Q56310; 115D.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0000155; P:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00065; GAF; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 683 AA; 77013 MW; 6989E95EB6543ED4 CRC64;

Query Match 81.0%; Score 34; DB 2; Length 683;
Best Local Similarity 88.9%; Pred. No. 83;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 548 QRIIASKSF 556

RESULT 9
Q678H2 PRELIMINARY; PRT; 310 AA.
AC Q678H2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Lymphocystis disease virus - isolate China.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Lymphocystivirus.
OX NCBI_TaxID=256729;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15194775; DOI=10.1128/JVI.78.13.6982-6994.2004;
Zhang Q.Y., Xiao F., Xie J., Li Z.Q., Gui J.F.;
"Complete Genome Sequence of Lymphocystis Disease Virus Isolated from
China.";
J. Virol. 78:6982-6994(2004).
DR EMBL; AY380826; AAU10885.1; --
DR InterPro; IPR001545; Gly_hormoneB.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.
DE Hypothetical protein.
SQ SEQUENCE 310 AA; 34349 MW; 714DAC2A33E42C11 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 310;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 QRIKASKSF 9
Db 183 QKIKSSKNF 191

RESULT 10
RLA0 THEAC STANDARD; PRT; 314 AA.
AC P57692;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Acidic ribosomal protein P0 homology (L10E).
GN Name=rlp0; OrderedLocusNames=Ta0359;
OS Thermoplasma acidophilum.
OC Archaea: Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
RA Ruepp A., Granel W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumelster W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513 (2000).
CC -!- FUNCTION: Ribosomal protein P0 is the functional equivalent of
CC E.coli protein L10.
CC -!- SIMILARITY: Belongs to the ribosomal protein L10P family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL445064; CAC11503.1; --
DR HANAP; MF 00280; ; 1.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00466; Ribosomal_L10; 1.
KW Complete proteome; Ribosomal protein.
SQ SEQUENCE 314 AA; 34583 MW; F40D048690E0B6EB CRC64;

Query Match 78.6%; Score 33; DB 1; Length 314;
Best Local Similarity 87.5%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKS 8
Db 17 QRIKASRS 24

RESULT 11
KAIN HUMAN STANDARD; PRT; 427 AA.
AC P29622; Q36B25;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kallistatin precursor (Kallikrein inhibitor) (Protease inhibitor 4).
GN Name=SRPINA4; Synonyms=kST, PI4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94043294; PubMed=8227002;
RA Chai K.X., Chen L.-M., Chao J., Chao L.;
RT "Kallistatin: a novel human serine proteinase inhibitor. Molecular

```

```

RT cloning, tissue distribution, and expression in Escherichia coli.";
RL J. Biol. Chem. 268:24498-24505 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95137583; PubMed=7835886;
RA Chai K.X., Ward D.C., Chao J., Chao L.;
RT "Molecular cloning, sequence analysis, and chromosomal localization of
RT the human protease inhibitor 4 (kallistatin) gene (PI4).";
RL Genomics 23:370-378 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausberg R.D., Colling F.S., Wagner L., Shenmen C.M., Schut G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 388-403.
RC TISSUE=Plasma;
RX MEDLINE=93100304; PubMed=1334488;
RA Zhou G.X., Chao L., Chao J.;
RT "Kallistatin: a novel human tissue kallikrein inhibitor. Purification,
RT characterization, and reactive center sequence.";
RL J. Biol. Chem. 267:25873-25880 (1992).
RN [5]
RP CARBOHYDRATE-LINKAGE SITE ASN-157.
RX MEDLINE=22660472; PubMed=12754519; DOI=10.1038/nbt827;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666 (2003).
CC -!- FUNCTION: Inhibits human amidolytic and kininogenase activities of
CC human tissue kallikrein. Inhibition is achieved by formation of an
CC equimolar, heat- and SDS-stable complex between the inhibitor and
CC the enzyme, and generation of a small C-terminal fragment of the
CC inhibitor due to cleavage at the reactive site by tissue
CC kallikrein.
CC -!- SUBUNIT: Monomer and some homodimers.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- PTM: The N-terminus is blocked.
CC -!- MISCELLANEOUS: Heparin blocks kallistatin's complex formation with
CC tissue kallikrein and abolishes its inhibitory effect on tissue
CC kallikrein's activity.
CC -!- SIMILARITY: Belongs to the serpin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19684; AAA59454.1; --
DR EMBL; L28101; AAC41706.1; --

```

DR EMBL; BC014992; AAH14992.1; --  
DR PIR; A49518; A49518.  
DR HSPF; P01011; 3CAA.  
DR Genew; HGNC:8948; SERPIN4.  
DR H-InvDB; HIX0011930; --  
DR MIM; 147935; --  
DR GO; GO:0005515; F:protein binding; TAS.  
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; TAS.  
DR InterPro; IPR000215; Prot\_inh\_serpin.  
DR Pfam; PF00079; Serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; 1.  
KW Direct protein sequencing; Glycoprotein; Plasma;  
KW Serine protease inhibitor; Serpin; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 427  
FT SITE 388 389 Reactive bond.  
FT CARBOHYD 33 33 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 108 108 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 157 157 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 238 238 N-linked (GlcNAc...) (Potential).  
FT CONFLICT 382 382 T -> S (in Ref. 3).  
SQ SEQUENCE 427 AA; 48556 MW; 3DBBE7AF956D4DAC CRC64;

Query Match 78.6%; Score 33; DB 1; Length 427;  
Best Local Similarity 66.7%; Pred. No. 85;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
|:::|||||  
Db 355 QKLEASKSF 363

RESULT 12  
Q86TR9 PRELIMINARY; PRT; 445 AA.  
AC Q86TR9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Full-length cDNA 5-PRIME end of clone CSODM009YC13 of Fetal liver of  
DE Homo sapiens (human) (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal liver;  
RA Li W.B., Gruber C., Jessee J., Polayes D.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal liver;  
RA Genoscope;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the serpin family.  
DR EMBL; BX248760; CAD66567.1; --  
DR HSPF; P01011; 3CAA.  
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.  
DR InterPro; IPR002086; Aldehyd dehydrog.  
DR InterPro; IPR000215; Prot\_inh\_serpin.  
DR Pfam; PF00079; Serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00070; ALDEHYDE DEHYDR\_CYS; UNKNOWN\_1.  
DR PROSITE; PS00284; SERPIN; 1.  
KW Protease inhibitor; Serine protease inhibitor; Serpin.  
FT NON TER 1  
SQ SEQUENCE 445 AA; 50612 MW; B1021CE4EC606079 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 445;  
Best Local Similarity 66.7%; Pred. No. 89;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
|:::|||||  
Db 373 QKLEASKSF 381

RESULT 13  
Q8EUP9 PRELIMINARY; PRT; 671 AA.  
AC Q8EUP9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Heavy-metal transporting P-type ATPase.  
GN OrderedLocusNames=MYPE8710;  
OS Mycoplasma penetrans.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=28227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HF-2;  
RX MEDLINE=22354719; PubMed=12466555; DOI=10.1093/nar/gkf667;  
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,  
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;  
RT "The complete genomic sequence of Mycoplasma penetrans, an  
RT intracellular bacterial pathogen in humans.";  
RL Nucleic Acids Res. 30:5293-5300(2002).  
DR EMBL; AP004173; BAC44663.1; --  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0015662; F:ATPase activity; coupled to transmembrane m...; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid...; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0030001; P:metal ion transport; IEA.  
DR InterPro; IPR006416; ATPase-IB hvy.  
DR InterPro; IPR001757; ATPase-El-E2.  
DR InterPro; IPR005834; Dehal\_like Hydro.  
DR InterPro; IPR008250; El-E2\_ATPase\_reg.  
DR Pfam; PF00122; El-E2\_ATPase; 1.  
DR Pfam; PF00702; Hydrolase; 1.  
DR PRINTS; PR00119; CATATPASE.  
DR TIGRFAMs; TIGR01525; ATPase-IB hvy; 1.  
DR TIGRFAMs; TIGR01494; ATPase-P-type; 2.  
DR PROSITE; PS00154; ATPASE\_El\_E2; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 671 AA; 75854 MW; 8EEE005CA3E60671 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 671;  
Best Local Similarity 77.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
|:::|||||  
Db 551 KEIKASKSF 559

RESULT 14  
RL10 MYCPE STANDARD; PRT; 161 AA.  
AC Q8EVJ0;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE 50S ribosomal protein L10.  
GN Name=rpLJ; OrderedLocusNames=MYPE5740;  
OS Mycoplasma penetrans.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=28227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HF-2;  
RX MEDLINE=22354719; PubMed=12466555; DOI=10.1093/nar/gkf667;

```
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
CC -!- SIMILARITY: Belongs to the ribosomal protein L10P family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP004172; BAC44364.1; ALT_INIT.
DR HAMAP; MF_00362; -; 1.
DR InterPro; IPR001790; Ribosomal L10.
DR InterPro; IPR002363; Ribosomal L10eub.
DR Pfam; PF00466; Ribosomal L10; 1.
DR PROSITE; PS01109; RIBOSOMAL_L10; FALSE_NEG.
KW Complete proteome; Ribosomal protein.
SQ SEQUENCE 161 AA; 17949 MW; 1D95D1C29B5E87D3 CRC64;
Query Match 76.2%; Score 32; DB 1; Length 161;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRIKASKSF 9
Db 17 EQIKSKSF 25
::||:||||
17 EQIKSKSF 25

RESULT 15
Q6FDR5 PRELIMINARY; PRT; 240 AA.
AC Q6FDR5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cell division inhibitor, inhibits ftsZ ring formation.
GN Name=minC; OrderedLocusNames=ACIAD0895;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP SEQUENCE FROM N.A.
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 0:0-0(2004).
DR EMBL; CR543861; CAG67793.1; -.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR InterPro; IPR005526; MinC.
DR Pfam; PF03775; MinC_C; 1.
DR TIGRFAMs; TIGR01222; minC; 1.
KW Complete proteome.
SQ SEQUENCE 240 AA; 26116 MW; F110BF30D94DFD90 CRC64;
Query Match 76.2%; Score 32; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRIKASK 7
Db 95 QRIKASK 101
|||||
95 QRIKASK 101
```

Search completed: July 13, 2005, 17:29:31  
Job time : 64.2222 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: July 13, 2005, 16:58:04 ; Search time 19.4444 Seconds  
(without alignments)  
34.552 Million cell updates/sec

Title: US-09-905-083A-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/1aa/5A COMB pep:\*
  - 2: /cgn2\_6/ptodata/1/1aa/5B COMB pep:\*
  - 3: /cgn2\_6/ptodata/1/1aa/6A COMB pep:\*
  - 4: /cgn2\_6/ptodata/1/1aa/6B COMB pep:\*
  - 5: /cgn2\_6/ptodata/1/1aa/PCTUS COMB pep:\*
  - 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	3	US-09-502-600-99
2	42	100.0	9	3	US-09-502-600-130
3	42	100.0	9	4	US-09-918-243-99
4	42	100.0	9	4	US-09-918-243-130
5	42	100.0	144	4	US-09-918-243-130
6	42	100.0	154	3	US-09-618-259-4
7	42	100.0	224	3	US-09-261-416-7
8	42	100.0	225	2	US-08-944-483-33
9	42	100.0	225	2	US-08-557-146-12
10	42	100.0	225	2	US-09-027-337-4
11	42	100.0	225	2	US-09-154-344-12
12	42	100.0	225	4	US-09-644-600-4
13	42	100.0	225	4	US-09-654-600A-4
14	42	100.0	253	2	US-08-557-146-2
15	42	100.0	253	2	US-08-824-874-3
16	42	100.0	253	2	US-09-154-344-2
17	42	100.0	253	3	US-08-930-188-2
18	42	100.0	253	4	US-09-210-084-3
19	42	100.0	253	4	US-09-764-762-3
20	42	100.0	265	4	PCT-US96-04294-2
21	36	85.7	9	3	US-09-949-016-7716
22	36	85.7	9	4	US-09-502-600-97
23	31	73.8	35	4	US-09-918-243-97
24	31	73.8	132	4	US-09-149-476-619
25	31	73.8	250	1	US-09-513-999C-5132
26	31	73.8	250	1	US-07-914-282D-6
27	31	73.8	250	5	US-08-276-887A-6
					PCT-US93-02460-6

28 31 73.8 314 4 US-09-248-796A-19386 Sequence 19386, A  
29 31 73.8 2317 4 US-09-949-016-10109 Sequence 10109, A  
30 31 73.8 2860 2 US-08-826-267-2 Sequence 2, Appli  
31 30 71.4 342 4 US-09-252-991A-18101 Sequence 18101, A  
32 30 71.4 666 3 US-08-982-785A-11 Sequence 11, Appl  
33 30 71.4 666 4 US-09-629-498-11 Sequence 11, Appl  
34 30 71.4 726 4 US-09-949-016-5916 Sequence 5916, Ap  
35 30 71.4 737 4 US-09-949-016-11607 Sequence 11607, A  
36 29 69.0 73 4 US-09-248-796A-26710 Sequence 26710, A  
37 29 69.0 203 4 US-09-710-279-2424 Sequence 2424, Ap  
38 29 69.0 222 3 US-09-134-001C-3400 Sequence 3400, Ap  
39 28 66.7 61 1 US-08-194-211A-3 Sequence 3, Appli  
40 28 66.7 61 3 US-08-456-748B-3 Sequence 3, Appli  
41 28 66.7 61 4 US-09-492-709A-357 Sequence 357, App  
42 28 66.7 67 4 US-09-489-039A-10090 Sequence 10090, A  
43 28 66.7 108 4 US-09-248-796A-17972 Sequence 17972, A  
44 28 66.7 159 4 US-09-134-000C-4565 Sequence 4565, Ap  
45 28 66.7 188 4 US-09-248-796A-27949 Sequence 27949, A

## ALIGNMENTS

### RESULT 1

US-09-502-600-99

; Sequence 99, Application US/09502600A

; Patent No. 6294344

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of

; FILE REFERENCE: D6223CIP-C

; CURRENT FILING DATE: 2000-02-11

; CURRENT APPLICATION NUMBER: US/09/502,600A

; PRIOR APPLICATION NUMBER: 09/039,211

; PRIOR FILING DATE: 03-14-1998

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 99

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Residues 91-99 of the SCCE protein

US-09-502-600-99

Query Match 100.0%; Score 42; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

DB 1 QRIKASKSF 9

### RESULT 2

US-09-502-600-130

; Sequence 130, Application US/09502600A

; Patent No. 6294344

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of

; FILE REFERENCE: D6223CIP-C

; CURRENT FILING DATE: 2000-02-11

; CURRENT APPLICATION NUMBER: US/09/502,600A

; PRIOR APPLICATION NUMBER: 09/039,211

; PRIOR FILING DATE: 03-14-1998

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 130

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-09-502-600-130

Query Match 100.0%; Score 42; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | |  
DB 1 QRIKASKSF 9

## RESULT 3

US-09-918-243-99  
; Sequence 99, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 99  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-09-918-243-99

Query Match 100.0%; Score 42; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | |  
DB 1 QRIKASKSF 9

## RESULT 4

US-09-918-243-130  
; Sequence 130, Application US/09918243  
; Patent No. 6627403  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 130  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-09-918-243-130

Query Match 100.0%; Score 42; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | |  
DB 1 QRIKASKSF 9

## RESULT 5

US-09-618-259-4  
; Sequence 4, Application US/09618259  
; Patent No. 6642013  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease  
; FILE REFERENCE: D6020CIP2  
; CURRENT APPLICATION NUMBER: US/09/618,259  
; CURRENT FILING DATE: 2000-07-18  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 72  
; SEQ ID NO 4  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: unknown  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic  
; OTHER INFORMATION: enzyme (scce) catalytic domain  
US-09-618-259-4

Query Match 100.0%; Score 42; DB 4; Length 144;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | |  
DB 28 QRIKASKSF 36

## RESULT 6

US-09-261-416-7  
; Sequence 7, Application US/09261416A  
; Patent No. 6291663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease  
; FILE REFERENCE: D6192  
; CURRENT APPLICATION NUMBER: US/09/261,416A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 14  
; SEQ ID NO 7  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum  
; OTHER INFORMATION: chymotryptic enzyme (scce) homologous to similar  
; OTHER INFORMATION: domain in TADG-12  
US-09-261-416-7

Query Match 100.0%; Score 42; DB 3; Length 154;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | |  
DB 37 QRIKASKSF 45

## RESULT 7

US-08-944-483-33  
; Sequence 33, Application US/08944483

; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; TITLE OF INVENTION: OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 224 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6232456e  
; US-08-944-483-33  
  
Query Match 100.0%; Score 42; DB 3; Length 224;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QRIKASKSF 9  
DB 62 QRIKASKSF 70  
  
RESULT 8  
US-08-557-146-12  
; Sequence 12, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York

; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sterner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 225 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
; US-08-557-146-12  
  
Query Match 100.0%; Score 42; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QRIKASKSF 9  
DB 63 QRIKASKSF 71  
  
RESULT 9  
US-09-027-337-4  
; Sequence 4, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotooshi  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in  
; Tissue of Invention: TADG-15: Breast and Ovarian Carcinomas  
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027,337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of Sccc homologous to  
; OTHER INFORMATION: similar domain in TADG-15  
; US-09-027-337-4  
  
Query Match 100.0%; Score 42; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QRIKASKSF 9  
DB 63 QRIKASKSF 71  
  
RESULT 10  
US-09-154-344-12  
; Sequence 12, Application US/09154344  
; Patent No. 5981256  
; GENERAL INFORMATION:

APPLICANT: Egelrud, Torbjorn  
APPLICANT: Hansson, Lennart  
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
TITLE OF INVENTION: Enzyme (SCCE)  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case, Patent Department  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,344  
FILING DATE: 16-SEP-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
US-09-154-344-12

Query Match 100.0%; Score 42; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
Db 63 QRIKASKSF 71

RESULT 11  
US-09-644-600-4  
Sequence 4, Application US/09644600  
Patent No. 6451500  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
APPLICANT: Tanimoto, Hirotooshi  
TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease  
FILE REFERENCE: D6064CIP/D  
CURRENT APPLICATION NUMBER: US/09/644,600  
CURRENT FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: 09/421,213  
PRIOR FILING DATE: 1999-10-20  
PRIOR APPLICATION NUMBER: 09/027,337  
PRIOR FILING DATE: 1998-02-20  
NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 4  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: SCCE

US-09-644-600-4  
Query Match 100.0%; Score 42; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QRIKASKSF 9  
Db 63 QRIKASKSF 71

RESULT 12  
US-09-654-600A-4  
Sequence 4, Application US/09654600A  
Patent No. 6649741  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
APPLICANT: Tanimoto, Hirotooshi  
TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease  
FILE REFERENCE: D6064CIP/D  
CURRENT APPLICATION NUMBER: US/09/654,600A  
CURRENT FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 09/421,213  
PRIOR FILING DATE: 1999-10-20  
1998-02-20  
NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 4  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: SCCE  
US-09-654-600A-4

Query Match 100.0%; Score 42; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
Db 63 QRIKASKSF 71

RESULT 13  
US-08-557-146-2  
Sequence 2, Application US/08557146  
Patent No. 5834290  
GENERAL INFORMATION:  
APPLICANT: Egelrud, Torbjorn  
APPLICANT: Hansson, Lennart  
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
TITLE OF INVENTION: Enzyme (SCCE)  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case, Patent Department  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.



/ REGISTRATION NUMBER: 35,372  
/ REFERENCE/DOCKET NUMBER: 1103326-181  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (212) 819-8783  
/ TELEFAX: (212) 354-8113  
/ INFORMATION FOR SEQ ID NO: 2:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 253 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-08-557-146-2

Query Match 100.0%; Score 42; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

RESULT 14  
US-08-824-874-3  
/ Sequence 3, Application US/08824874  
/ Patent No. 5962300  
/ GENERAL INFORMATION:  
/ APPLICANT: Hillman, Jennifer L.  
/ APPLICANT: Lal, Preeti  
/ TITLE OF INVENTION: NOVEL KALLIKREIN  
/ NUMBER OF SEQUENCES: 5  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.  
/ STREET: 3174 Porter Drive  
/ CITY: Palo Alto  
/ STATE: CA  
/ COUNTRY: USA  
/ ZIP: 94304  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Diskette  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: DOS  
/ SOFTWARE: Fast-SEQ for Windows Version 2.0  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/824,874  
/ FILING DATE: Filed Herewith  
/ CLASSIFICATION: 514  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER:  
/ FILING DATE:  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Billings, Lucy J.  
/ REGISTRATION NUMBER: 36,749  
/ REFERENCE/DOCKET NUMBER: PF-0252 US  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 415-855-0555  
/ TELEFAX: 415-845-4166  
/ INFORMATION FOR SEQ ID NO: 3:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 253 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ IMMEDIATE SOURCE:  
/ LIBRARY: GenBank  
/ CLONE: 532504  
US-08-824-874-3

Query Match 100.0%; Score 42; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

RESULT 15  
US-09-154-344-2  
/ Sequence 2, Application US/09154344  
/ Patent No. 5981256  
/ GENERAL INFORMATION:  
/ APPLICANT: Egelrud, Torbjorn  
/ APPLICANT: Hansson, Lennart  
/ TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
/ TITLE OF INVENTION: Enzyme (SCCE)  
/ NUMBER OF SEQUENCES: 17  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: White & Case, Patent Department  
/ STREET: 1155 Avenue of the Americas  
/ CITY: New York  
/ STATE: New York  
/ COUNTRY: U.S.A.  
/ ZIP: 10036-2787  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/154,344  
/ FILING DATE: 16-SEP-1998  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/557,146  
/ FILING DATE: 14-DEC-1995  
/ CLASSIFICATION:  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Sterner, Richard J.  
/ REGISTRATION NUMBER: 35,372  
/ REFERENCE/DOCKET NUMBER: 1103326-181  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (212) 819-8783  
/ TELEFAX: (212) 354-8113  
/ INFORMATION FOR SEQ ID NO: 2:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 253 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-09-154-344-2

Query Match 100.0%; Score 42; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

Search completed: July 13, 2005, 17:34:30  
Job time : 20.4444 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 13, 2005, 17:29:45 ; Search time 64.8889 Seconds  
(without alignments)  
53.584 Million cell updates/sec

Title: US-09-905-083A-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 1726220

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	42	100.0	9	9 US-09-918-243-99	Sequence 99, Appl
2	42	100.0	9	9 US-09-918-243-130	Sequence 130, Appl
3	42	100.0	9	9 US-09-905-083-99	Sequence 99, Appl
4	42	100.0	9	9 US-09-905-083-130	Sequence 130, Appl
5	42	100.0	9	15 US-10-372-521-99	Sequence 99, Appl
6	42	100.0	9	15 US-10-372-521-130	Sequence 130, Appl
7	42	100.0	9	16 US-10-831-075-99	Sequence 99, Appl
8	42	100.0	9	16 US-10-831-075-130	Sequence 130, Appl
9	42	100.0	97	15 US-10-262-511-100	Sequence 100, Appl
10	42	100.0	144	9 US-09-796-294-4	Sequence 4, Appli
11	42	100.0	144	14 US-10-461-787-4	Sequence 4, Appli

12	42	100.0	181	15	US-10-262-511-98	Sequence 98, Appl
13	42	100.0	198	15	US-10-262-511-96	Sequence 96, Appl
14	42	100.0	224	11	US-09-789-210-33	Sequence 33, Appl
15	42	100.0	224	15	US-10-262-511-104	Sequence 104, Appl
16	42	100.0	225	15	US-10-600-187-4	Sequence 4, Appli
17	42	100.0	247	15	US-10-262-511-102	Sequence 102, Appl
18	42	100.0	250	15	US-10-262-511-92	Sequence 92, Appl
19	42	100.0	252	15	US-10-262-511-94	Sequence 94, Appl
20	42	100.0	253	9	US-09-888-615-98	Sequence 98, Appl
21	42	100.0	253	9	US-09-764-762-3	Sequence 3, Appli
22	42	100.0	253	14	US-10-071-214-2	Sequence 2, Appli
23	42	100.0	253	14	US-10-071-214-48	Sequence 48, Appl
24	42	100.0	253	14	US-10-264-283-90	Sequence 90, Appl
25	42	100.0	253	15	US-10-295-027-498	Sequence 498, Appl
26	42	100.0	253	15	US-10-173-999-48	Sequence 48, Appl
27	42	100.0	253	16	US-10-408-765A-639	Sequence 639, Appl
28	42	100.0	253	16	US-10-643-795A-95	Sequence 95, Appl
29	42	100.0	253	17	US-10-948-518-95	Sequence 95, Appl
30	42	100.0	253	17	US-10-868-490A-1	Sequence 1, Appli
31	42	100.0	257	15	US-10-344-394-38	Sequence 38, Appl
32	39	92.9	226	14	US-10-071-214-49	Sequence 49, Appl
33	36	85.7	9	9	US-09-918-243-97	Sequence 97, Appl
34	36	85.7	9	15	US-09-905-083-97	Sequence 97, Appl
35	36	85.7	9	16	US-10-831-075-97	Sequence 97, Appl
36	36	85.7	249	14	US-10-071-214-50	Sequence 50, Appl
37	36	85.7	243	14	US-10-210-172-122	Sequence 46, Appl
38	33	78.6	411	15	US-10-210-172-122	Sequence 122, Appl
39	33	78.6	411	15	US-10-210-172-124	Sequence 124, Appl
40	33	78.6	411	15	US-10-210-172-126	Sequence 126, Appl
41	33	78.6	411	15	US-10-210-172-128	Sequence 128, Appl
42	33	78.6	411	15	US-10-210-172-130	Sequence 130, Appl
43	33	78.6	411	15	US-10-210-172-132	Sequence 132, Appl
44	33	78.6	411	15	US-10-210-172-134	Sequence 134, Appl
45	33	78.6	411	15	US-10-210-172-134	Sequence 134, Appl

## ALIGNMENTS

RESULT 1  
US-09-918-243-99  
; Sequence 99, Application US/09918243  
; Patent No. US20020142317A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP  
; CURRENT APPLICATION NUMBER: US/09/918,243  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 136  
; SEQ ID NO 99  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-09-918-243-99

Query Match 100.0%; Score 42; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
Db 1 QRIKASKSF 9

RESULT 2

```
US-09-918-243-130
; Sequence 130, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 130
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-918-243-130

Query Match          100.0%; Score 42; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 3
US-09-905-083-99
; Sequence 99, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/DiV
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 99
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-905-083-99

Query Match          100.0%; Score 42; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 4
US-09-905-083-130
; Sequence 130, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/DiV
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 130
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-905-083-130

Query Match          100.0%; Score 42; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 5
US-10-372-521-99
; Sequence 99, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 99
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-10-372-521-99

Query Match          100.0%; Score 42; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 1 QRIKASKSF 9

RESULT 6
US-10-372-521-130
; Sequence 130, Application US/10372521
; Publication No. US20030223973A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP2
; CURRENT APPLICATION NUMBER: US/10/372,521
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/918,243
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 130
; LENGTH: 9
```

TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CHAIN  
OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-10-372-521-130

Query Match 100.0%; Score 42; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | | |  
DB 1 QRIKASKSF 9

## RESULT 7

US-10-831-075-99  
; Sequence 99, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 99  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-10-831-075-99

Query Match 100.0%; Score 42; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | | |  
DB 1 QRIKASKSF 9

## RESULT 8

US-10-831-075-130  
; Sequence 130, Application US/10831075  
; Publication No. US20040224891A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Cannon, Martin J.  
; APPLICANT: Santin, Alessandro  
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer  
; FILE REFERENCE: D6223CIP/C/D/CIP3  
; CURRENT APPLICATION NUMBER: US/10/831,075  
; CURRENT FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 10/372,521  
; PRIOR FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 130  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: Residues 91-99 of the SCCE protein  
US-10-831-075-130

Query Match 100.0%; Score 42; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
| | | | | | | | | |  
DB 1 QRIKASKSF 9

## RESULT 9

US-10-262-511-100  
; Sequence 100, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 100  
; LENGTH: 97  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-10-262-511-100

Query Match 100.0%; Score 42; DB 15; Length 97;  
Best Local Similarity 100.0%; Pred. No. 0.69; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
|||||  
Db 72 QRIKASKSF 80

## RESULT 10

US-09-796-294-4  
; Sequence 4, Application US/09796294  
; Patent No. US20020037581A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Extracellular Serine Protease  
; FILE REFERENCE: D6020CIP3  
; CURRENT APPLICATION NUMBER: US/09/796,294  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 09/618,259  
; PRIOR FILING DATE: 2000-07-18  
; NUMBER OF SEQ ID NOS: 72  
; SEQ ID NO 4  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: unknown  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic  
; OTHER INFORMATION: enzyme (scce) catalytic domain  
US-09-796-294-4

Query Match 100.0%; Score 42; DB 9; Length 144;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
|||||  
Db 28 QRIKASKSF 36

## RESULT 11

US-10-461-787-4  
; Sequence 4, Application US/10461787  
; Publication No. US20030199010A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: No. US20030199010A1el Extracellular Serine Protease  
; FILE REFERENCE: D6020CIP2  
; CURRENT APPLICATION NUMBER: US/10/461,787  
; CURRENT FILING DATE: 2003-06-13  
; PRIOR APPLICATION NUMBER: US/09/618,259  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: US 09/127,444  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 72  
; SEQ ID NO 4  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: unknown  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic  
; OTHER INFORMATION: enzyme (scce) catalytic domain  
US-10-461-787-4

Query Match 100.0%; Score 42; DB 14; Length 144;  
Best Local Similarity 100.0%; Pred. No. 1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
|||||  
Db 28 QRIKASKSF 36

## RESULT 12

US-10-262-511-98  
; Sequence 98, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 98  
; LENGTH: 181  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-262-511-98

Query Match 100.0%; Score 42; DB 15; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 1.3; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

QY 1 QRIKASKSF 9  
 Db 72 QRIKASKSF 80

RESULT 13  
 US-10-262-511-96

; Sequence 96, Application US/10262511  
 ; Publication No. US20040038223A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Smithson, Glennda  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Peyman, John A.  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Ju, Jingfang  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Guo, Xiaojia (Sasha)  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Ort, Tatiana  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Zerhusen, Bryan D.  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Pena, Carol E. A.  
 ; APPLICANT: Shenoy, Suresh G.  
 ; APPLICANT: Shimkets, Richard A.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Berghs, Constance  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-462C  
 ; CURRENT APPLICATION NUMBER: US/10/262,511  
 ; PRIOR FILING DATE: 2003-05-28  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/373,815  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,642  
 ; PRIOR FILING DATE: 2002-05-17  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,038  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/373,260  
 ; PRIOR FILING DATE: 2002-04-17  
 ; PRIOR APPLICATION NUMBER: 60/373,826  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 60/327,435  
 ; PRIOR FILING DATE: 2001-10-05  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 96  
 ; LENGTH: 198

; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-262-511-96

Query Match 100.0%; Score 42; DB 15; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 1.4; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

QY 1 QRIKASKSF 9  
 Db 91 QRIKASKSF 99

RESULT 14

US-09-789-210-33  
 ; Sequence 33, Application US/09789210  
 ; Publication No. US20040241646A1

; GENERAL INFORMATION:

; APPLICANT: COHEN, MAURICE  
 ; COLPITTS, TRACEY L.  
 ; FRIEDMAN, PAULA N.  
 ; GRANADOS, EDWARD N.  
 ; KLASS, MICHAEL R.  
 ; RUSSELL, JOHN C.  
 ; STEWART, KENT D.  
 ; STROUPE, STEVEN D.

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
 ; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
 ; OF THE PROSTATE

; NUMBER OF SEQUENCES: 76

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories  
 ; STREET: 100 Abbott Park Road  
 ; CITY: Abbott Park  
 ; STATE: IL  
 ; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/789,210

; FILING DATE: 20-Feb-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/944,483

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Becker, Cheryl L.

; REGISTRATION NUMBER: 35,441

; REFERENCE/DOCKET NUMBER: 6183.US.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 847/935-1729

; TELEFAX: 847/938-2623

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 224 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: None

; SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Query Match 100.0%; Score 42; DB 11; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9  
 |||||

Db 62 QRIKASKSF 70

```
RESULT 15
US-10-262-511-104
; Sequence 104, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Saeha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 104
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-104
```

Query Match 100.0%; Score 42; DB 15; Length 224;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9  
| | | | | | | | | |  
Db 77 QRIKASKSF 85

Search completed: July 13, 2005, 18:23:42  
Job time : 64.8889 secs